

https://dspace.org/

EEstimating children's weight in the emergency department using a length-based, habitus-modified system to increase the accuracy of total body weight estimation: the utility of the PAWPER tape in paediatric emergency medicine ta

Wells, Mike

2018

https://hdl.handle.net/10539/27968

Wells, Mike (2018) Estimating children's weight in the emergency department using a length - based, habitus-modified system to increase the accuracy of total body weight estimation:the utility of the PAWPER tape in paediatric emergency medicine, University of the Witwatersrand, Johannesburg, <http://hdl.handle.net/10539/27968>

Downloaded from a DSpace repository. https://dspace.org/

Estimating children's weight in the Emergency Department using a length-based, habitus-modified system to increase the accuracy of total body weight estimation: the utility of the PAWPER tape in paediatric emergency medicine

Mike Wells

A thesis submitted to the Faculty of Health Sciences, University of the Witwatersrand, in fulfilment of the requirements for the degree of Doctor of Philosophy.

Johannesburg, 2018

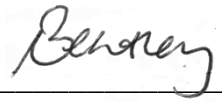
Declaration

Candidate's contribution to articles and agreement of co-authors

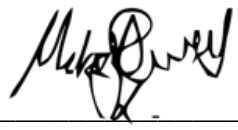
I, Michael David John Wells, student number 8523767, declare that this thesis is my own work and that I contributed adequately towards the research findings published in the articles stated below, which are included in my thesis. I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong. I confirm that ALL the work submitted is my own unaided work except where I have explicitly indicated otherwise. I have followed the required conventions in referencing the thoughts and ideas of others. I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

Supervisor – Dr Alison Bentley

Signature of Supervisor



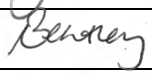
Agreement by co-authors: by signing this declaration, the co-authors listed below agree to the use of the article as part of this thesis.



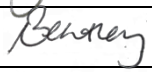
(Signature of candidate)

29th day of November 2018 in Johannesburg

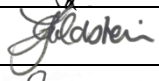
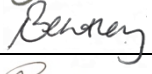
Wells M, Goldstein L, Bentley A. The accuracy of emergency weight estimation systems in children - a systematic review and meta-analysis. International Journal of Emergency Medicine 2017;10(29):1-43. https://doi.org/10.1186/s12245-017-0156-5

Author	**Contribution**
Mike Wells	***Percentage contribution: 80%.*** *Specific roles: formulation and development of research goals and aims. Development and design of the methodology or created models. Performance of oversight and leadership responsibility for the research activity planning and execution. Data collection and management including the development of software code, to derive or analyse data. Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data. Preparation, creation and presentation of the published work, specifically data presentation. Preparation, creation and critical review and revision of the manuscript.*
Lara Goldstein	***Percentage contribution: 10%.*** *Specific roles: assistance in the development and design of the methodology. Facilitation and support of the research activity planning and execution. Data collection. Critical review and revision of the manuscript.*
Alison Bentley	***Percentage contribution: 10%.*** *Specific roles: mentorship on formulation and development of research goals and aims, development and design of the methodology. Critical review, commentary on the manuscript.*
1st Author	
2nd Author	
3rd Author	

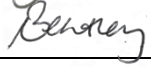
Wells M, Goldstein L, Bentley A. It is time to abandon age-based emergency weight estimation in children! A failed validation of 20 different age-based formulas. Resuscitation 2017;116(7):73-83 http://dx.doi.org/10.1016/j.resuscitation.2017.05.018

Author	**Contribution**
Mike Wells	***Percentage contribution: 80%.*** *Specific roles: formulation and development of research goals and aims. Development and design of the methodology or created models. Performance of oversight and leadership responsibility for the research activity planning and execution. Data collection and management including the development of software code, to derive or analyse data. Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data. Preparation, creation and presentation of the published work, specifically data presentation. Preparation, creation and critical review and revision of the manuscript.*
Lara Goldstein	***Percentage contribution: 10%.*** *Specific roles: assistance in the development and design of the methodology. Facilitation and support of the research activity planning and execution. Data collection. Critical review and revision of the manuscript.*
Alison Bentley	***Percentage contribution: 10%.*** *Specific roles: mentorship on formulation and development of research goals and aims, development and design of the methodology. Critical review, commentary on the manuscript.*
1st Author	
2nd Author	
3rd Author	

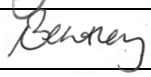
Wells M, Goldstein L, Bentley A, Basnett S, Monteith I. The accuracy of the Broselow tape as a weight estimation tool and a drug-dosing guide - a systematic review and meta-analysis. Resuscitation 2017;121:9-17. http://dx.doi.org/10.1016/j.resuscitation.2017.09.026.

Author	Contribution
Mike Wells	***Percentage contribution: 80%.*** *Specific roles: formulation and development of research goals and aims. Development and design of the methodology or created models. Performance of oversight and leadership responsibility for the research activity planning and execution. Data collection and management including the development of software code, to derive or analyse data. Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data. Preparation, creation and presentation of the published work, specifically data presentation. Preparation, creation and critical review and revision of the manuscript.*
Lara Goldstein	***Percentage contribution: 10%.*** *Specific roles: assistance in the development and design of the methodology. Facilitation and support of the research activity planning and execution. Data collection. Critical review and revision of the manuscript.*
Alison Bentley	***Percentage contribution: 10%.*** *Specific roles: mentorship on formulation and development of research goals and aims, development and design of the methodology. Critical review, commentary on the manuscript.*
Sian Basnett	***Percentage contribution: 5%.*** *Specific roles: assistance in the research activity execution. Data collection. Critical review of the manuscript.*
Iain Monteith	***Percentage contribution: 5%.*** *Specific roles: assistance in the research activity execution. Data collection. Critical review of the manuscript.*
1st Author	
2nd Author	
3rd Author	
4th Author	
5th Author	

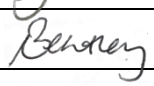
Wells M, Goldstein L, Bentley A. A validation study of the PAWPER XL tape: accurate estimation of both total and ideal body weight in children up to 16 years of age. Trauma and Emergency Care 2017;2(4):1-8 http://dx.doi.org/10.15761/TEC.1000141

Author	**Contribution**
Mike Wells	***Percentage contribution: 80%.*** *Specific roles: formulation and development of research goals and aims. Development and design of the methodology or created models. Performance of oversight and leadership responsibility for the research activity planning and execution. Data collection and management including the development of software code, to derive or analyse data. Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data. Preparation, creation and presentation of the published work, specifically data presentation. Preparation, creation and critical review and revision of the manuscript.*
Lara Goldstein	***Percentage contribution: 10%.*** *Specific roles: assistance in the development and design of the methodology. Facilitation and support of the research activity planning and execution. Data collection. Critical review and revision of the manuscript.*
Alison Bentley	***Percentage contribution: 10%.*** *Specific roles: mentorship on formulation and development of research goals and aims, development and design of the methodology. Critical review, commentary on the manuscript.*
1st Author	
2nd Author	
3rd Author	

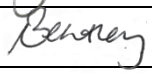
Wells M, Goldstein L, Bentley A. The use of body habitus reference images improves the ability of novices to accurately estimate children's weight using the PAWPER XL tape system. The Journal of Emergency Medicine 2017;54(2):165-175. https://doi.org/10.1016/j.jemermed.2017.10.009

Author	**Contribution**
Mike Wells	***Percentage contribution: 80%.*** *Specific roles: formulation and development of research goals and aims. Development and design of the methodology or created models. Performance of oversight and leadership responsibility for the research activity planning and execution. Data collection and management including the development of software code, to derive or analyse data. Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data. Preparation, creation and presentation of the published work, specifically data presentation. Preparation, creation and critical review and revision of the manuscript.*
Lara Goldstein	***Percentage contribution: 10%.*** *Specific roles: assistance in the development and design of the methodology. Facilitation and support of the research activity planning and execution. Data collection. Critical review and revision of the manuscript.*
Alison Bentley	***Percentage contribution: 10%.*** *Specific roles: mentorship on formulation and development of research goals and aims, development and design of the methodology. Critical review, commentary on the manuscript.*
1st Author	
2nd Author	
3rd Author	

Wells M, Goldstein L, Bentley A. Development and validation of a method to estimate bodyweight in critically ill children using length and mid-arm circumference measurements - the PAWPER XL-MAC system. South African Medical Journal 2017;107(11):1015-1021. https://doi.org/10.7196/SAMJ.2017.v107i11.12505

Author	**Contribution**
Mike Wells	***Percentage contribution: 80%.*** *Specific roles: formulation and development of research goals and aims. Development and design of the methodology or created models. Performance of oversight and leadership responsibility for the research activity planning and execution. Data collection and management including the development of software code, to derive or analyse data. Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data. Preparation, creation and presentation of the published work, specifically data presentation. Preparation, creation and critical review and revision of the manuscript.*
Lara Goldstein	***Percentage contribution: 10%.*** *Specific roles: assistance in the development and design of the methodology. Facilitation and support of the research activity planning and execution. Data collection. Critical review and revision of the manuscript.*
Alison Bentley	***Percentage contribution: 10%.*** *Specific roles: mentorship on formulation and development of research goals and aims, development and design of the methodology. Critical review, commentary on the manuscript.*
1st Author	
2nd Author	
3rd Author	

Wells M, Goldstein L, Bentley A. The accuracy of paediatric weight estimation during simulated emergencies: the effects of patient position, patient cooperation and human errors. African Journal of Emergency Medicine 2018;8(2):43-50. https://doi.org/10.1016/j.afjem.2017.12.003.

Author	**Contribution**
Mike Wells	***Percentage contribution: 80%.*** *Specific roles: formulation and development of research goals and aims. Development and design of the methodology or created models. Performance of oversight and leadership responsibility for the research activity planning and execution. Data collection and management including the development of software code, to derive or analyse data. Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data. Preparation, creation and presentation of the published work, specifically data presentation. Preparation, creation and critical review and revision of the manuscript.*
Lara Goldstein	***Percentage contribution: 10%.*** *Specific roles: assistance in the development and design of the methodology. Facilitation and support of the research activity planning and execution. Data collection. Critical review and revision of the manuscript.*
Alison Bentley	***Percentage contribution: 10%.*** *Specific roles: mentorship on formulation and development of research goals and aims, development and design of the methodology. Critical review, commentary on the manuscript.*
1st Author	
2nd Author	
3rd Author	

Dedication

This thesis is dedicated to my amazing wife Lara, without whom it would never have been completed.

List of presentations

International conferences

Conference	International Conference on Emergency Medicine
Title of presentation	The PAWPER tape: a new length-based, habitus-modified tape system to estimate children's weight in emergencies
Date	July 2012
Location	Dublin, Ireland

Conference	World Conference on Disaster and Emergency Medicine
Title of presentation	Weight estimation in medical emergencies in children: a meta–analysis of the accuracy of current methodologies.
Date	22 April 2015
Location	Cape Town, South Africa

Local conferences

Conference	Emergency Care Society of South Africa conference
Title of presentation	Children: a vulnerable population
Date	18 September 2014
Location	Maropeng, South Africa

Conference	"*Up to spaed*" paediatric symposium
Title of presentation	Fat is the new normal
Date	2 August 2015
Location	Johannesburg, South Africa

Conference	University of the Witwatersrand, Faculty of Health Sciences research day
Title of presentation	Weight estimation in medical emergencies in children: a meta-analysis of the accuracy of current methodologies.
Date	30 September 2015
Location	Johannesburg, South Africa

Conference	"*Up to spaed*" paediatric symposium
Title of presentation	An update on the PAWPER tape.
Date	28 July 2017
Location	Johannesburg, South Africa

Publications

List of publications arising from this work

1 Wells M, Goldstein L, Bentley A. The accuracy of emergency weight estimation systems in children - a systematic review and meta-analysis. International Journal of Emergency Medicine 2017;10(29):1-43. https://doi.org/10.1186/s12245-017-0156-5

2 Wells M, Goldstein L, Bentley A. A systematic review and meta-analysis of the accuracy of weight estimation systems used in paediatric emergency care in developing countries. African Journal of Emergency Medicine 2017;7:S36-S54. https://doi.org/10.1016/j.afjem.2017.06.001

3 Wells M, Goldstein L, Bentley A. It is time to abandon age-based emergency weight estimation in children! A failed validation of 20 different age-based formulas. Resuscitation 2017;116(7):73-83. http://dx.doi.org/10.1016/j.resuscitation.2017.05.018

4 Wells M, Goldstein L, Basnett S, Monteith I, Bentley A. The accuracy of the Broselow tape as a weight estimation tool and a drug-dosing guide – a systematic review and meta-analysis. Resuscitation 2017;121:9-17. http://dx.doi.org/10.1016/j.resuscitation.2017.09.026

5 Wells M, Goldstein L, Bentley A. A validation study of the PAWPER XL tape: accurate estimation of both total and ideal body weight in children up to 16 years of age. Trauma and Emergency Care 2017;2(4):1-8. http://dx.doi.org/10.15761/TEC.1000141

6 Wells M, Goldstein LN. Optimising emergency weight estimation in underweight and obese children: the accuracy of estimations of total body weight and ideal body weight
European Journal of Emergency Medicine 2018. In Press. http://dx.doi.org/10.1097/MEJ.0000000000000561

7 Wells M, Goldstein L, Bentley A. The use of body habitus reference images improves the ability of novices to accurately estimate children's weight using the PAWPER XL tape system. The Journal of Emergency Medicine 2017;54(2):165-175. https://doi.org/10.1016/j.jemermed.2017.10.009

8 Wells M, Goldstein L, Bentley A. Accuracy of weight estimation by the Broselow tape is substantially improved by including a modifier based on a visual assessment of body habitus. Pediatric Research 2018;83(1):83-92. http://dx.doi.org/10.1038/pr.2017.222

9 Wells M, Goldstein L, Bentley A. Development and validation of a method to estimate bodyweight in critically ill children using length and mid-arm circumference measurements - the PAWPER XL-MAC system. South African Medical Journal 2017;107(11):1015-1021. https://doi.org/10.7196/SAMJ.2017.v107i11.12505

10 Wells M, Goldstein L, Bentley A. High-tech adjuncts to emergency weight estimation: Point-of-care ultrasound and point-of-care bioelectrical impedance measurements can increase the accuracy of length-based weight estimation in children. Trauma and Emergency Care 2017;2(3):1-8. http://dx.doi.org/10.15761/TEC.1000134

11 Wells M, Goldstein L, Bentley A. The accuracy of paediatric weight estimation during simulated emergencies: the effects of patient position, patient cooperation and human errors. African Journal of Emergency Medicine 2018;8(2):43-50. https://doi.org/10.1016/j.afjem.2017.12.003

12 Wells M, Goldstein LN. An algorithm to improve the accuracy of emergency weight estimation in obese children. Pan-African Medical Journal 2018;31:90. https://doi.org/10.11604/pamj.2018.31.90.13821

Abstract

An accurate estimation of weight during the management of medical emergencies in children is essential to permit effective intervention and to avoid medication errors. Although older methods of weight estimation, such as age-based formulas and the Broselow tape, are still being used, they have failed to achieve an acceptable standard of accuracy. The newer, dual length- and habitus-based systems, such as the PAWPER tape system, have been evaluated in a wide variety of populations and shown to be significantly more accurate. However, some limitations in the PAWPER system have been identified including the subjectivity of habitus assessment and a relatively lower weight estimation accuracy in very obese children. The increasing prevalence of severe obesity in both first and third world countries requires a strategy to mitigate the increased risk of poor outcomes from drug dosing errors in obese children. This requires an increased sophistication from weight estimation systems as both total body weight and ideal body weight are required for appropriate drug dosing. The series of studies that make up this thesis was designed to develop an evidence base from which to refine and validate the second version of the PAWPER tape (the PAWPER XL tape). It was also intended to further clarify its optimum design as the imperative for accurate and appropriate weight estimation cannot be ignored. Since any increased complexity in weight estimation methodology may lead to heightened risk of errors, the upgraded PAWPER XL system was specifically formulated to minimise the potential for errors through developing cognitively neutral weight estimation processes, embracing technological adjuvants and advocating for adequate and appropriate training in weight estimation techniques.

Acknowledgements

Most importantly, I would like to gratefully acknowledge the support and assistance of my co-researcher and wife, Professor Lara Goldstein. There is no doubt that without her this research and this thesis would never have been completed. I would also like to express my gratitude to my supervisor, Dr Alison Bentley, for her assistance.

Finally, I would like to express my sincere thanks to the following companies:

Vertec South Africa	who provided the Discovery Hologic A DXA machine on loan.
I Sandler & Co.	who provided the TANITA SC240MA on loan.
Drs Monteith & partners	who provided partial sponsorship for the development of the EDD4C app.
Fyrebird creations	who produced the images for the figural reference image system (Liesel Roux).

Table of contents

		Page
Declaration		ii
Dedication		x
List of presentations		xi
International conferences		xi
Local conferences		xii
List of publications		xiii
Abstract		xv
Acknowledgements		xvi
Table of contents		xvii
Nomenclature and abbreviations		xx
Nomenclature		xx
Abbreviations		xxii
List of tables and figures		xxiv
Section 1 introductory chapters		1
Chapter 1	Organisation of the thesis	2
Chapter 2	Introduction	5
	2.1 Preamble	7
	2.2 The accuracy of weight estimation systems	11
	2.3 The usability of weight estimation methods	11
	2.4 Integration of weight estimation and drug dosing	13
	Literature survey	14
	2.5 The impact of weight estimation errors	16
	2.6 Weight estimation accuracy standards	18
	2.7 The performance of weight estimation systems	22
	2.8 Two-dimensional weight estimation systems	29
	2.9 What is the best way of assessing children's body habitus?	42
	2.10 Can anthropometry be used to assess habitus?	43
	2.11 How do cognitive demands affect weight estimation?	44

2.12 Should weight estimation be modified in obese children? 45
2.13 Can mobile phone apps improve weight estimation? 49
2.14 How do real-world factors cause weight estimation errors? 50
2.15 Summary 51
2.16 Gaps in the literature 52
Motivation for the thesis 56
2.17 Problem statements 57
Chapter 3 Aims and objectives 83
Section 2 The accuracy of weight estimation systems 84
Chapter 4 The accuracy of emergency weight estimation systems in children - a systematic review and meta-analysis. 85
Chapter 5 It is time to abandon age-based emergency weight estimation in children! A failed validation of 20 different age-based formulas. 147
Chapter 6 The accuracy of the Broselow tape as a weight estimation tool and a drug-dosing guide – a systematic review and meta-analysis. 169
Section 3 Refining the function of the PAWPER tape system 208
Chapter 7 A validation study of the PAWPER XL tape: accurate estimation of both total and ideal body weight in children up to 16 years of age. 209
Chapter 8 The use of body habitus reference images improves the ability of novices to accurately estimate children's weight using the PAWPER XL tape system. 225
Chapter 9 Development and validation of a method to estimate bodyweight in critically ill children using length and mid-arm circumference measurements - the PAWPER XL-MAC system. 248
Chapter 10 The accuracy of paediatric weight estimation using the Broselow tape, the Mercy method and the PAWPER XL tape during simulated emergencies: the effects of patient position, patient cooperation and human errors. 267
Section 4 Final discussion and conclusions 287
Chapter 11 Integrated discussion 288
11.1 Introduction 290

11.2 The accuracy of weight estimation systems 291
11.3 Refining the performance of the PAWPER system 298
11.4 The usability of weight estimation systems 306
11.5 Integration of weight estimation and drug dosing 312
11.6 Teaching and training 316
11.7 The need for mobile phone technology 318
11.8 Summary of key findings 319
Chapter 12 Integrated conclusion 333

Appendixes 340
Appendix 1 PhD proposal I
Appendix 2 Ethics clearance documents XXVII
Appendix 3 Plagiarism screening report XXX

Nomenclature and abbreviations

Nomenclature

Accuracy	The closeness of estimates to the exact or true values that the methods were intended to measure. For weight estimation studies, this is the most important statistic, from a statistical and pragmatic point of view. Accuracy is most often reported as the percentage of weight estimates that are within 10% or 20% of the actual measured weight (PW10 or PW20).
Bias or trueness	The tendency of a weight-estimation system to over- or underestimate actual weight. This is most often represented by mean error or mean deviation, or mean percentage error (MPE).
Dose scalar	The base weight format that is used to convert a mg/kg dose to a mg dose. It may be total body weight (TBW), ideal body weight (IBW), lean body weight (LBW) or an adjusted body weight (ABW).
Drug dosing	The process whereby a drug dose is determined and calculated.
Habitus score	The numerical score assigned to a child which represents their habitus (or degree of adiposity). The score runs from 1 to 5 for the PAWPER tape and from 1 to 7 for the PAWPER XL tape with an HS3 representing the “average” child, HS1 and HS2 thin children and HS4, HS5, HS6 and HS7 overweight, obese and severely obese children.
Normal weight child	A child who is not underweight, overweight or obese.
Obese	A child who has a weight-for-age or BMI-for-age greater than the 95th centile (Z-score of 2).
One-dimensional weight estimation system	A system that makes use of only a single variable to predict weight such as age, or length.
Overweight	A child who has a weight-for-age or BMI-for-age between the 90th and 95th centiles.
Pharmacokinetic mass	The theoretical ideal mass to which an individual drug should be scaled in an individual child, to provide optimum pharmacokinetics.
Precision	This refers to a weight-estimation system being highly reproducible or having a small dispersion of estimates, although these estimates might not necessarily be close to the true or actual weight. This may be represented as limits

	of agreement (LOA), percentage error limits of agreement (PELOA), root mean squared error (RMSE) or root mean squared percentage error (RMSPE).
Resuscitation	This refers to the management of any acute, potentially life-threatening medical or traumatic event, not only cardiac arrest. Conditions include, *inter alia*, major trauma, status epilepticus, status asthmaticus, severe sepsis, cardiac arrest.
Severely obese	A child who has a weight-for-age or BMI-for-age greater than the 97th centile (Z-score of 2.5).
Two-dimensional weight estimation system	A system that makes use of two variables to predict weight. Also referred to as dual length- and habitus-based weight estimation as the systems make use of length (or a surrogate measurement for length) and habitus (or a surrogate for habitus) to generate a weight prediction.
Underweight	A child who has a weight-for-age or BMI-for-age less than the 5thth centile (Z-score of -2).
Weight descriptor	As for *Dose scalar*.

Abbreviations

1D	One-Dimensional
2D	Two-Dimensional
3D	Three-Dimensional
APLS	Advanced Paediatric Life Support (Formula)
ARC	Australian Resuscitation Council (Formula)
BG	Best Guess (Formula)
BMI	Body Mass Index (kg/m^2)
BT	Broselow Tape
C	Correlation
CAWR	Chinese Age-Weight Rule (Formula)
DWEM	Devised Weight-Estimating Method
ED	Emergency Department
EMS	Emergency Medical Services
EPLS	European Life Support (Formula)
FE	Fixed Effects
FM	Fat Mass
FMI	Fat Mass Index
FFM	Fat-Free Mass
FFMI	Fat-Free Mass Index
HCP	Healthcare Provider
IBW	Ideal Body Weight
ICC	Inter-Class Correlation
IQR	Interquartile Range
IV	Inverse Variance
LBW	Lean Body Weight
LOA	95% Limits of Agreement (Kg)
LOE	Level of Evidence
MAC	Mid-Arm Circumference
MD	Mean Difference (Kg)

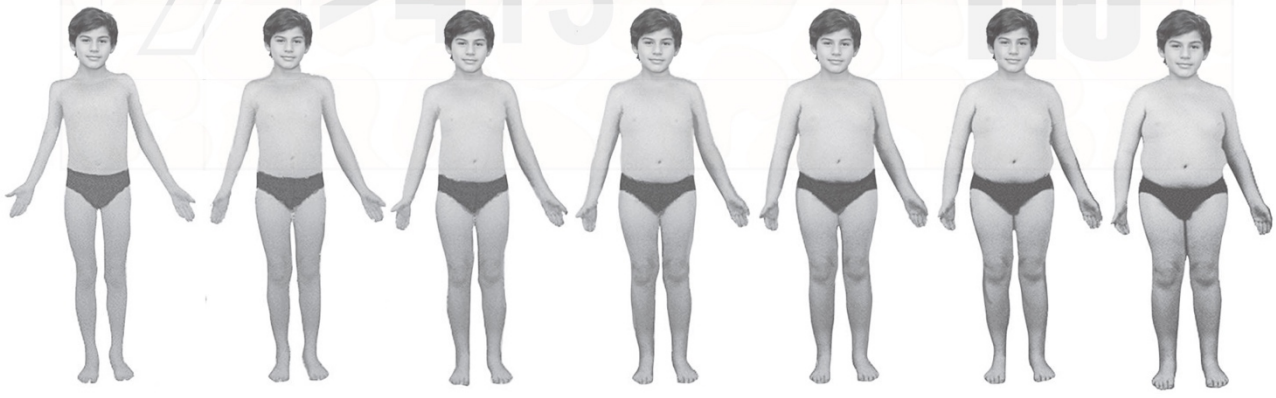
Mo	Months
MPE	Mean Percentage Error
NHANES	National Health and Nutrition Examination Survey
OR	Odds Ratio
NCHS	National Centre for Health Statistics
PAWPER	Paediatric Advanced Weight Prediction in the Emergency Room
PELOA	Percentage Error 95% Limits of Agreement (%)
PICU	Paediatric Intensive Care Unit
PW10	Percentage of estimates Within 10% of Actual Weight
PW20	Percentage of estimates Within 20% of Actual Weight
PW30	Percentage of estimates Within 30% of Actual Weight
R	Regression analysis
RE	Random Effects
RMSE	Root Mean Squared Error (kg)
RMSPE	Root Mean Squared Percentage Error
TBW	Total Body Weight
TJ	Traub-Johnson Formula
TK	Traub-Kichen Formula
Wks	Weeks
Wt	Weight (Kg)
Yrs	Years

List of tables and figures

Table 1-1	Contents of the chapters in Section 1	2
Table 1-2	Contents of the chapters in Section 2	3
Table 1-3	Contents of the chapters in Section 3	4
Table 1-4	Contents of the chapters in Section 4	4
Table 1-5	Contents of the Appendixes	5
Figure 2-1	A flow diagram illustrating the sequence of events involved in prescribing and administering drugs in an emergency.	9
Table 2-1	Characteristics of weight estimation systems that will influence the accuracy of the system.	12
Figure 2-2	Timeline of significant events in the history of weight estimation.	15
Figure 2-3	The finger-counting method	24
Figure 2-4	The Broselow tape	26
Figure 2-5	The red zone of the Broselow tape.	27
Figure 2-6	The Mercy table.	31
Figure 2-7	The Mercy 3D tape.	32
Figure 2-8	The PAWPER tape.	33
Figure 2-9	A sample segment of the PAWPER tape.	34
Figure 2-10	The Flipper is a low-cost drug dosing guide for children.	35
Figure 2-11	The relative positions of the tape weight divisions for the Broselow tape and PAWPER tape.	36
Figure 2-12	Position of the tape at the child's head.	37
Figure 2-13	Position of the tape at the child's heel.	37
Figure 2-14	The correct measurement technique for the PAWPER tape – side view.	37
Figure 2-15	The correct measurement technique for the PAWPER tape – top view.	37
Figure 2-16	A length segment of the PAWPER tape showing the HS3 weight.	38
Figure 2-17	A length segment of the PAWPER tape showing the other HS weights.	38
Figure 2-18	A cropped image of the cover of the Emergency Drug Dosing in Children book cover.	39
Figure 2-19	The Collins figural images for boys.	43
Figure 2-20	Weight descriptors to use for underweight, normal weight and obese children.	48
Figure 2-21	The black box model adapted to cognitive aids.	50

Suppl Table 4-1	Data from each study included in the meta-analysis, including subgroup data where available.	133
Suppl Figure 4-1	The full Review manager (Revman 5.3) figures for each direct comparison between weight estimation systems.	143
Suppl Figure 5-1	Figural reference images for habitus score assessment.	150
Suppl Table 6-1	Modified Newcastle-Ottawa scale as used in this study.	199
Suppl Table 6-2	Raw data from the included studies.	200
Suppl Figure 8-1	Figural reference images with frontal and lateral views for three age-groups.	241
Suppl Table 9-1	The performance data of the PAWPER XL-MAC system in the derivation and validation stages, shown segment-by-segment.	259
Suppl Table 9-2	The results of the bootstrap analysis performed on the NHANES data.	261
Figure 10-1	Screenshots from the EDD4C app.	268
Suppl Table 10-1	Representative responses and key themes obtained from the post-participation survey.	281
Table 10-1	The influence of each factor on the accuracy of the weight estimation method.	282
Table 11-1	Changes to the PAWPER tape system.	299
Figure 11-1	A summary of the major findings from this research.	321
Table 12-1	The classification of recommendations and level of evidence used in this chapter.	336
Table 12-2	Potential future research questions.	337

Section 1
Introductory chapters



Chapter 1

Organisation of the thesis

There are four Sections in this thesis, plus the Appendixes. The contents of the chapters within each section are described below. The chapters which contain the individual research articles (chapters 4 to 10) have a prologue which provides some background and context for the article and an epilogue which links the findings to the next chapter. The references for the prologue and epilogue are included at the end of the epilogue for convenience.

1.1 Section 1

Introductory chapters

This section contains the preliminary chapters. The details are tabulated Table 1-1.

Table 1-1 Contents of the chapters in Section 1.

Chapter 1	**Organisation of the thesis.**
Chapter 2	**Introduction.** This chapter provides contextual information on why paediatric weight estimation is important and about some of the weight estimation systems discussed in the thesis. The major gaps which exist in the literature relating to emergency weight estimation and emergency drug dosing are highlighted and a brief survey of the relevant literature is included. The chapter concludes with the problem statements and the motivation for the thesis.
Chapter 3	**Aims and objectives.**

1.2 Section 2

The accuracy of weight estimation systems

The evidence concerning the accuracy of existing weight estimation systems is presented in this section. It contains two meta-analyses: the first meta-analysis evaluated the accuracy of the existing traditional (one-dimensional) and newer dual length- and habitus-based (two-dimensional) systems; and the second was aimed at evaluating the accuracy of the Broselow tape as a weight estimation instrument and

drug-dosing guide. This section also contains a study which comprehensively evaluated and questioned the use of age-based weight estimation techniques. The details are shown in Table 1-2.

Table 1-2 Contents of the chapters in Section 2.

Chapter 4	A systematic review and meta-analysis on the accuracy of weight estimation systems.
Chapter 5	A study on the accuracy of age-based weight estimation formulas.
Chapter 6	A systematic review and meta-analysis on the Broselow tape.

1.3 Section 3

Refining the function of the PAWPER tape system

The evidence used to upgrade and validate the performance of the PAWPER XL tape system is described in this section. It contains a study validating the PAWPER XL tape and evaluating the ability of the tape to predict IBW, which is (ideally) required for drug dose calculations in obese children. A study comparing three methods of assessing children's habitus was used to determine the most effective and accurate method of assessing habitus that should be used with the PAWPER XL system. There is a study which describes the development and validation of a fully objective, anthropometric method of PAWPER XL tape weight estimation, the PAWPER XL-MAC method. The final section contains a description of the evaluation of the EDD4C weight-estimation app in a simulation-based study. In this study, the accuracy and functioning of several weight estimation systems was evaluated using realistic, suboptimal patient positioning and cooperation. The chapter details are shown in Table 1-3.

Table 1-3 Contents of the chapters in Section 3.

Chapter 7	A validation study on the PAWPER XL tape.
Chapter 8	A study on different methods of habitus assessment for the PAWPER XL tape.
Chapter 9	The development and validation of the PAWPER XL-MAC system using mid-arm circumference to categorise habitus.
Chapter 10	A study on the accuracy and usability of different weight estimation during simulated paediatric emergencies.

1.4 Section 4

Final discussion, conclusions and recommendations

The final discussion, recommendations and suggested future research and overall conclusions chapters complete the thesis. The details are shown in Table 1-4.

Table 1-4 Contents of the chapters in Section 4.

Chapter 11	**Final discussion.** This section summarises the findings of the individual studies and relates them to the original motivation for the thesis.
Chapter 12	**Final conclusions.** This chapter summarises the major findings of the research as they relate to the initial aims and objectives. It also provides some recommendations which translate the findings of the individual studies into guidelines on improving patient safety with regards to weight estimation and drug dosing i.e. how we can care for children better. It also contains suggestions on what future research is needed in this field.

1.5 Appendixes

Table 1-5 shows the contents of the Appendixes.

Table 1-5 Contents of the Appendixes.

Appendix 1	The approved PhD research proposal and protocols.
Appendix 2	The ethics clearance certificates for the PhD studies.
Appendix 3	The plagiarism software report from Turnitin.

Chapter 2

Introduction

Contents

Introduction 5

2.1 Preamble 7

2.1.2 Why is weight estimation needed? 8

2.2 The accuracy of weight estimation systems 11

2.3 The usability of weight estimation methods 11

2.4 Integration of weight estimation and drug dosing 13

Literature survey 14

2.5 The impact of weight estimation errors 16

2.6 Weight estimation accuracy standards 18

2.6.1 How accurate should weight estimation systems be? 18

2.6.2 Current reference methods of emergency weight estimation 20

2.7 The performance of weight estimation systems 22

2.7.1 Overview 22

2.7.2 Age-based weight estimation 23

2.7.3 Length-based weight estimation 25

2.7.4 The Broselow tape 26

2.7.5 Mid-arm circumference-based weight estimation 27

2.7.6 Guesses of weight by healthcare providers 28

2.7.7 Estimates of weight by parents 28

2.7.8 Miscellaneous weight-estimation methods 28

2.8 Two-dimensional weight estimation systems 29

2.8.1 The Mercy method 30

2.8.2 The PAWPER tape system 32

2.9 What is the best way of assessing children's body habitus? 42

2.10 Can anthropometry be used to assess habitus? 43

2.11 How do cognitive demands affect weight estimation? 44

2.12 Should weight estimation be modified in obese children? 45

2.13 Can mobile phone applications improve weight estimation? 49

2.14 How do real-world factors cause weight estimation errors? 50

2.15 Summary 51

2.16 Gaps in the literature 52
Motivation for the thesis 56
2.17 Problem statements 57
References 61

2.1 Preamble

Paediatric weight estimation and medication safety are inextricably intertwined [1]. However, despite the many advancements in the science of resuscitation in children, not enough is known about the impact of weight estimation on medication safety. An accurate weight is clearly one of the key determinants of patient safety, as weight-based dosing is the norm, both during routine care and emergency care [2]. One of the confounding factors in appraising weight estimation methodology is that not enough is known about some of the fundamental aspects of paediatric pharmacology [3]. Even questions on the appropriate therapeutic doses of potentially life-saving medications are unresolved. Nonetheless, it is still essential to attempt to determine the impact of errors in weight estimation and drug dose preparation on outcomes and how special groups such as underweight or obese children should be managed in this context [4-6].

While paediatric emergency weight estimation is a very small part of the overall process of paediatric emergency care, it is, nevertheless, absolutely essential to ultimately ensure the safety of children receiving emergency treatment [7].

This chapter provides a brief overview of some of the important topics and concepts that will be presented in the chapters containing the research papers. Firstly, the importance of emergency weight estimation is discussed. Thereafter, the three main themes or concepts that are central to weight estimation research are introduced: matters relating to the accuracy of weight estimation systems, to the usability of weight estimation systems and to the integration of weight estimation systems with drug dosing guides. The research in this thesis was focused primarily on the accuracy of weight estimation methods, including the PAWPER tape systems, but also on the usability of these methods during emergency care. The integration of weight estimation and drug dosing systems was only considered from a theoretical perspective, in terms of how a weight estimation system should be designed in order to optimise this relationship.

2.1.2 Why is weight estimation needed?

Since virtually all drug doses are based on body weight in children, it is essential to have a quantification of weight in order to determine appropriate, safe and effective doses of oral and intravenous medications [8]. It would be ideal to have an accurate measurement of a child's weight at the time when medication is prescribed, but this is sometimes not possible. This may be because a scale is not available or because the child cannot be weighed because of the need to initiate emergency medical care [9, 10]. Some weight estimation methods have been proposed for any occasion when a properly calibrated scale is not available, especially in resource-limited settings [11]. In the context of this thesis, however, the focus will be on emergency weight estimation, whether it be required in the prehospital environment, the Emergency Department or the Paediatric Intensive Care Unit.

During an emergency, weight can seldom be measured accurately, even if sophisticated equipment such as in-stretcher scales are available [12]. Furthermore, while parents are frequently able to provide an accurate estimation of their child's weight, a parent might not be available at the time that emergency treatment needs to be initiated or might not be able to provide or recall a recently measured weight [13]. Under these circumstances an estimation of weight, that is sufficiently accurate to prevent medication errors, is required [14]. Without an accurate measurement or estimation of weight, accurate drug dose determination and delivery is not possible, which may directly result in increased morbidity or mortality [15-17]. The risk for a poor outcome is thus at least partly determined by the accuracy of the weight estimation and the accurate delivery of the correct dose of medication [18].

Figure 2-1 illustrates the process of weight estimation and drug dose calculations during paediatric emergency care.

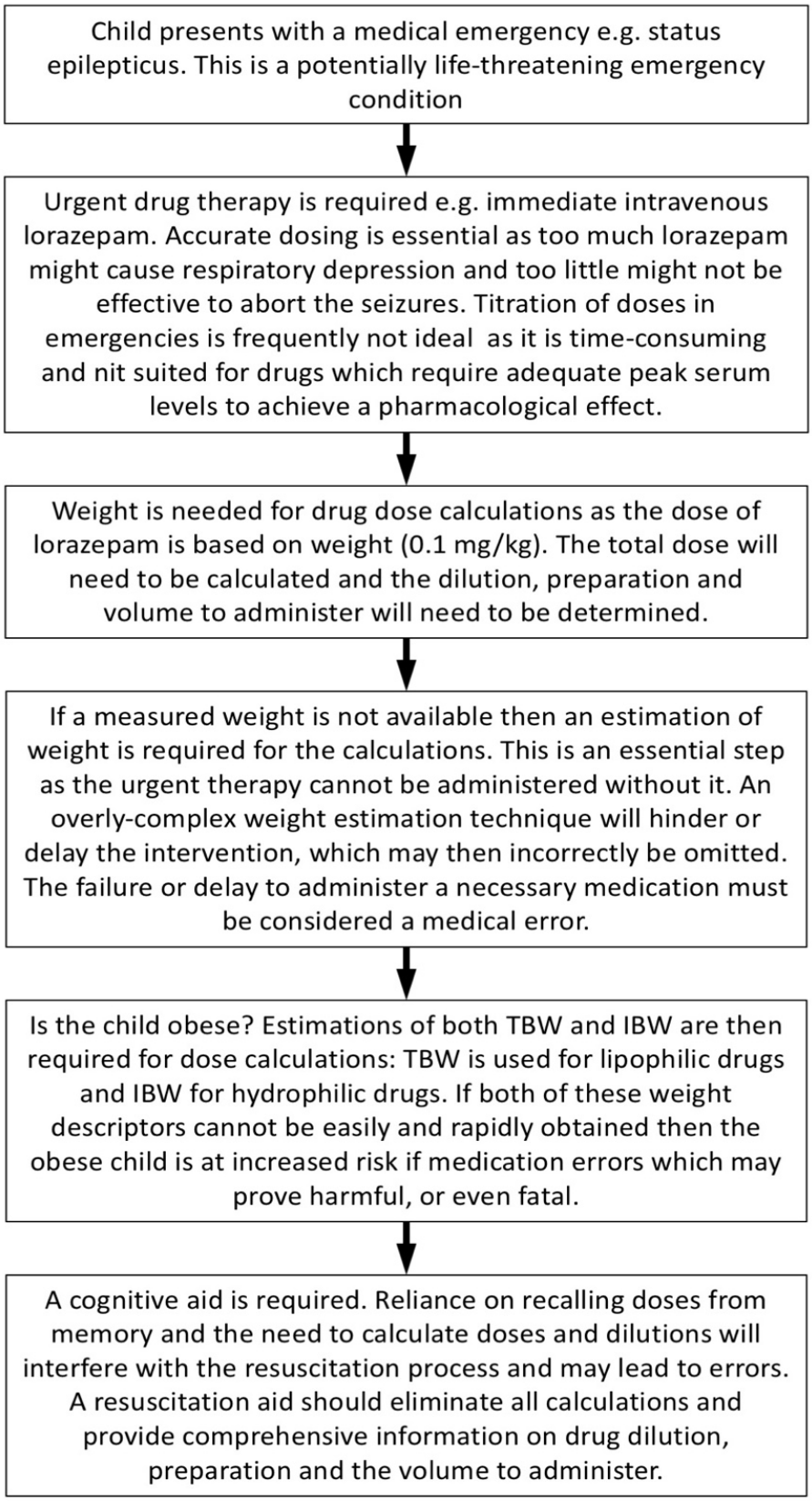


Figure 2-1 A flow diagram illustrating the sequence of events involved in prescribing and administering drugs in an emergency.

There is no high-quality evidence to indicate how often emergency weight estimation might be required in any Emergency Department. It is likely to vary from hospital to hospital, depending on the number of children seen at the hospital and the acuity of the patients. The best way to estimate how often emergency weight estimation might be needed, would be to look at the percentage of critically ill or injured children attending an Emergency Department, who are triaged as “critical” and require resuscitation. Studies from paediatric Emergency Departments in Asia, Europe and South Africa have shown that critical patients account for between 0.4% and 2.3% of all paediatric presentations (the highest figure was incidentally from South Africa) [19-22]. Perhaps of greater relevance, is that this proportion was equivalent to at least one and perhaps even two critical patients presenting every day. Although not common, children suffering from conditions that require resuscitation, and therefore emergency weight estimation, are seen daily. This affirms the importance and relevance of emergency weight estimation.

The relative rarity of critical paediatric cases, while not diminishing the need for emergency weight estimation tools, does create another problem related to weight estimation: the skills required for procedure that is not used often will decay rapidly over time [23]. This means that weight estimation methods need to not only be accurate, but also simple and usable enough for healthcare providers to retain the necessary skills to allow accurate use. Appropriate training and frequent practice is one of the key elements required to achieve this [24].

The performance of a weight estimation system under ideal circumstances may be very different to its performance during emergencies. Both its intrinsic accuracy as well as its usability are crucial factors that impact on its overall functioning. These concepts will be discussed below.

2.2 The accuracy of weight estimation systems

Alarmingly, the medication error rate in children has been shown to be as much as three times the rate found in adult patients. This is largely due to the need for weight-based dosing and the associated weight estimation errors [4]. Accurate weight estimation is, therefore, an important prerequisite for appropriate error-free paediatric drug dosing. There are a number of aspects of the performance of weight estimation systems that need to be considered in this regard:

- What is an acceptable standard of accuracy for a weight estimation system?
- Which weight estimation systems are the most accurate under ideal clinical circumstances?
- Which weight estimation systems are the most accurate under realistic emergency resuscitation circumstances?

This data would be useful to evaluate the appropriateness of each weight estimation system for specific clinical environments. The answers to these questions form a central element of this thesis and will be discussed in greater detail in the literature survey and the chapters with the individual research studies.

The definition of accuracy that will be used throughout this thesis is as follows (a detailed explanation will be presented in Chapter 4):

- A weight estimate can be considered accurate if it is within 10% of the child's actual measured weight.
- A weight estimation system can be considered accurate in a study if it predicts 70% of children's weights within 10% of their actual measured weight.

2.3 The usability of weight estimation methods

Weight estimation systems should be considered to be resuscitation aids and should be judged according to the standards applied to other cognitive aids [25]. This is important because the functional accuracy of the weight estimation system will be determined by these characteristics. The final accuracy of the drug dose administered will also be dependent on these traits, however, which is the most important consideration. The qualities that should be considered are shown in Table 2-1. These factors will determine the overall usability of a system.

Table 2-1 Characteristics of weight estimation systems that will influence the accuracy and usability of the system.

Characteristic	Description
Completeness	This refers to the system being autonomous with no need for additional components to function. It also refers to the ability of the system to minimise the cognitive load by eliminating the need for calculations and non-automatic thinking. *Complete systems* – Broselow tape, PAWPER tape. *Incomplete systems* – age formulas, Mercy method.
Simplicity (ease-of-use)	This refers to the subjective rating by users of the difficulty of using a system and the reliability of the required measurements. *Simple systems* – Broselow tape, PAWPER tape *Complex systems* – Mercy method, age formulas
Resilience	This refers to a system's ability to perform accurately under suboptimal or adverse conditions. No system is completely resilient, but the simpler systems have a higher degree of resilience than the other systems. Methods which rely on memory and calculations are the most vulnerable to error during emergencies.
Performance reinforcement	This refers to the ability of the system to prevent downstream errors i.e. a system that can provide complete information on weight, drug doses and equipment sizes. There are no such systems at present. Even the Broselow tape does not contain sufficient information on the tape about drug dose preparation to provide sufficient performance reinforcement to eliminate errors.

Usability is a critical consideration in the evaluation, development and adoption of current and future methods of emergency weight estimation [26]. Many of the

component elements of usability described in Table 2-1 will be evaluated in the studies contained in this thesis. The ideal weight estimation system must not only be accurate, it must be self-contained, easy-to-use, resilient and integrated with a comprehensive drug dosing guide. There are few published studies on this topic, but the scarce information that is available suggests that usability is an important additional determinant of the accuracy of a weight estimation system.

The concept of usability that will be used throughout this thesis is thus a measure of ease-of-use (accurate weight estimation despite adverse human factors such as lack of user training or experience) and resilience of the weight estimation system under typical clinical conditions (accurate weight estimation despite adverse patient factors such as suboptimal patient position and limited patient cooperation).

2.4 Integration of weight estimation and drug dosing

Accurate and usable weight estimation systems are facilitators of accurate drug dosing. The process that occurs between obtaining an estimation of weight and injecting a quantum of a potentially life-saving medication into a child may make a difference in the accuracy of the dose, even if the weight estimation is correct. This will not be covered in any depth in this thesis, as it is beyond the scope of the main themes, but it is important to understand the concepts, because they influence the development and functioning of weight estimation systems.

The only system that would maintain the downstream effects of accurate weight estimation perfectly would be a computerised system (or an app) which provided both weight estimation and drug dosing information. The user would only need to enter the appropriate variables (e.g. length and habitus score for the PAWPER tape or humeral length and mid-arm circumference for the Mercy method) and the output could include weight, drug doses and equipment sizes. The program would also be able to use TBW or IBW appropriately for dosing in obese children with no further input from the user.

While a system such as this would be essential in all populations, it might not be available in poorer communities. A less expensive option would then be a weight

estimation system with a companion drug dosing guide. As mentioned in Table 2-1, while the Broselow tape, and other length-based tapes, do have some information printed on the tape which can eliminate dose calculation errors, this information does not assist with the calculations that may cause preparation and administration errors.

An understanding of this concept is helpful in deciding on the best design for a weight estimation tape. It is also helpful to allow a user to critically evaluate the best weight estimation methods and the best drug dosing guides.

These concepts will be explored further in the chapters that follow.

Literature survey

This literature survey is not intended to provide a detailed literature review, but rather to provide a contextual outline of the relevant topics as they apply to the individual studies presented in the thesis. A focused literature review is presented for each of the research papers in their individual chapters. The status of knowledge at the beginning of the research for this thesis has been presented, unless otherwise stated.

The earliest record of the prediction of children's weight was in 1956, with the publication of an expected weight for age formula [27]. It was only in the 1980s that the next publications on weight estimation appeared. The two decades between 1988 and 2012 were dominated by the research on age- and length-based weight estimation. The period since 2012 has been dominated by the newer two-dimensional length- and habitus-based systems. A timeline of the development of weight estimation systems is shown in Figure 2-2.

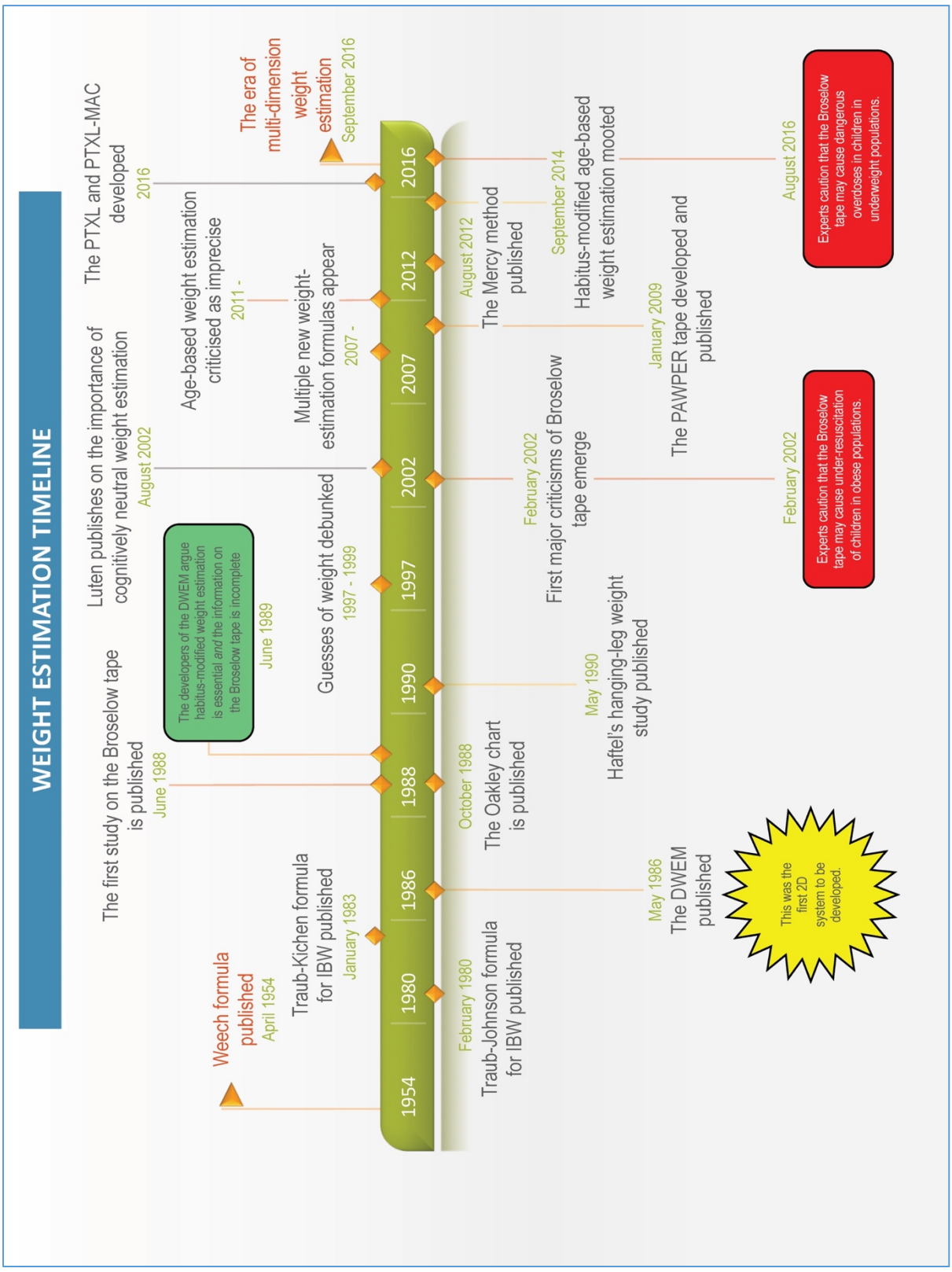


Figure 2-2 Timeline of significant events in the history of weight estimation.

2.5 The impact of weight estimation errors

Children are not safe in our Emergency Departments: medication errors, and especially dosing errors, are commonplace in paediatric emergency medicine. This has caused a great deal of concern, especially with the lack in progress in dealing with these errors in the last decade [28-30]. The American Academy of Pediatrics released a policy in 2011 entitled “Principles of pediatric patient safety: reducing harm due to medical care” in which they emphasised the importance of the establishment of a culture of safety, leadership, the use of information technology, the generation of patient-safety goals, the use of decision-support systems and the elimination of medication errors (and especially dosing errors) [31]. These principles are even more important during emergency care, where the stressful circumstances of the paediatric resuscitation create an environment that is highly conducive to error [32-36]. There is a growing concern that, in children, mistakes in the process of medication prescription and administration are more common than in adults, and are more significant, or have a higher potential for harm, than was previously realised [37, 38].

There is strong consensus that the most important intervention to reduce medication errors in any aspect of paediatric care is to ensure that adequate systems are in place to prevent and detect errors [39-42]. The need for such systems is even more apparent during emergency care and the eradication, or at least minimisation, of the potential root-cause errors in these circumstances is imperative [1, 43]. It is obvious that error-free emergency drug dosing is dependent on being able to neutralise each of the potential errors in the medication-administration process. The sources of error include inaccurate estimations of weight, errors in the calculation or administration of correct drug doses and cognitive overload (caused by insufficient training and over-reliance on memory together with under-reliance on technology).

The weight estimation process has been identified as a significant contributor medication errors in children, especially during emergency care [43, 44]. These errors are suspected to be a major source of morbidity and mortality in children undergoing resuscitation [15]. Not only are these errors potentially harmful, but they are highly prevalent with some studies having shown the medication error rate to be

more than 70% during resuscitation [18]. This is therefore a significant and serious problem. Specific risks related to paediatric weight-based dosing in the Emergency Department include not using the appropriate weight as well as making incorrect calculations [4]. It may be difficult to distinguish between weight estimation errors and calculation errors, especially during prehospital care, as both errors may occur together [45-47]. This highlights the need for a weight estimation system that integrates seamlessly with a resuscitation aid to eliminate as many sources of error as possible [18, 43].

When evaluating poor outcomes (paediatric patient morbidity), 20 to 30% of serious, preventable medication errors are as a result of not having essential patient information, mostly an accurate weight, at the time of medication administration [7, 17]. This number may be as high as 40% when less serious medication errors are considered [48]. When a weight estimation error of >10% is present, it leads to a significant medication error in more than 30% of cases [7]. The frequency and severity of errors is also directly correlated with severity of illness, with the sickest children at the greatest risk of error [2].

Weight estimation in obese children has been linked to poor outcomes. A higher mortality in obese children with cardiac arrest, when compared to normal weight children, has been speculated to be as a result of drug dosing errors caused by poor weight estimation [49]. Most currently available weight estimation methods are very inaccurate in overweight and obese children, which has been shown to result in large drug dosing errors [50].

While it might be difficult to separate weight estimation errors from cognitive and calculation errors during emergency care in the prehospital environment or in the Emergency Department it is apparent that these errors are frequent and potentially severe. It is therefore essential to eliminate as many sources of error as possible, beginning with using the best available weight estimation techniques [1].

Accurate estimation of weight is an essential element of the process to eliminate dosing errors incurred during paediatric emergencies.

2.6 Weight estimation accuracy standards

There are two aspects that need to be considered with regards to weight estimation accuracy standards: the benchmarks that are used for assessing the performance of weight estimation systems; and the weight estimation methods that are considered to be the reference “gold standards”.

2.6.1 How accurate should weight estimation systems be?

This topic is one of the important concepts addressed in this thesis. It is dealt with in greater detail in later chapters. The appraisal of a weight estimation error in an individual child is different to the assessment of a weight estimation system. The question for the weight estimation system must ultimately be: in what proportion of children in a population can an inaccurate weight estimate and therefore a medication error be tolerated?

One of the definitions of accuracy is, “deviating only slightly, or within acceptable limits from a standard” [51]. Implicit within this definition is the unequivocal requirement for a standard by which to judge the efficacy of a measurement system. When considering weight estimation systems in paediatric emergency care, there has only been one study in which a statistically appropriate standard of weight estimation accuracy has been proposed. This research report from Australia, evaluating the performance of the Broselow tape, provided a recommended minimum accuracy for a weight estimation system: at least 95% of weight estimations should be within 20% of actual, measured weight (PW20>95%). This equates to a critical error rate of less than 5% [52].

Other studies have suggested that tolerable errors of weight estimation should be based on what errors in drug dosing accuracy would be considered to be within acceptable limits (accurate to within 10%, 20% or even 30% of actual weight), but none have provided a benchmark by which to judge a weight-estimation system rather than an individual weight estimate [53, 54].

Although there is limited agreement on what deviation from a drug dose range would constitute a dose error, in the end, drug doses must still not be too high so as to

cause unwanted side-effects and not be too low so that the potentially life-saving efficacy is not achieved. There is little data on how accurate drug doses need to be to comply with these requirements [55]. Some authors suggest that a dose error of 30% or even 40% is acceptable [55, 56]. Unfortunately, there is no satisfactory, evidence-based answer to this question [57, 58].

It is clear that the required dosing accuracy is different for different drugs and in different clinical circumstances. The drugs used during emergencies generally have a higher risk for toxic effects than, for example, most antibiotics. A 30% overdose of penicillin would not have the same toxic consequences as the same overdose of amiodarone. Similarly, a 30% underdosing of penicillin would not have quite the same lack-of-efficacy consequences as there might be with underdosing lorazepam during status epilepticus [59]. Overall, many authorities would agree that a dose deviation of 10% is probably tolerable, but that a 20% deviation should be regarded as a dosing error [15, 42, 60]. New evidence has shown that when a weight estimation error of >10% is incurred, a significant, potentially harmful medication error results in more than 30% of cases [7].

Critical dosing errors with narrow therapeutic window drugs produce a high risk of poor patient outcomes and, from an ethical standpoint, should not be accepted. In acknowledging the lack of strong evidence to guide what proportion of errors is acceptable, two options are possible:

- to accept that it does not matter what we do, because we do not know if it makes any difference [57]; or
- to attempt to be as accurate as is pragmatically achievable and to eliminate as many sources of errors as possible [61].

Adopting an approach to minimise errors as much as possible is more ethically appropriate, until better evidence emerges.

Furthermore, the contribution of other errors in the drug administration process, during dilution and preparation, may add substantially to the cumulative error [1]. A weight estimation error of 20% may escalate to a dose error of 50% with

compounded errors, resulting in a much higher potential for patient harm [62]. In order to minimise potential injury, or lack of efficacy, from dosing errors the highest accuracy achievable with a practical (low cognitive load) methodology should be adopted.

2.6.2 Current reference methods of emergency weight estimation

The gold standard for children's weight for drug dosing calculations is an accurate, measured weight. This is seldom available during emergency care [63]. Although the use of actual weight as a factor to calculate drug doses seems logical, the manufacturers of the Broselow tape dispute this [55]:

> "Considering the wide therapeutic range of most medications, there is currently no evidence that for most resuscitation medications, giving a dose based on actual body weight is superior to giving an IBW dose and titrating to clinical effect if necessary. Indeed, using the actual weight could increase the risk of toxicity in certain clinical situations."

Although this statement from 10 years ago is outdated, the belief illustrates the complexity of the circumstances. Current thinking is that TBW should be used to calculate all drug doses for all non-obese children. In obese children, both TBW *and* IBW may be required to appropriately dose different medications, although there is currently no evidence on whether such complex and intricate calculations would be practical during emergency care [5, 64]. Furthermore, the practice of titrating medications is possible for some drugs but not for others [61].

Since the major advanced life support organisations in the USA and UK (the American Heart Association and the European Resuscitation Council) provide guidelines and standards for emergency and resuscitative care, it would be helpful if these organisations provided standards for emergency weight estimation. Unfortunately, the European Resuscitation Council provides no guidance on paediatric weight estimation in its publications. The American Heart Association does have some published some guidelines [65, 66]:

> "If the child's weight is unknown, it is reasonable to use a body length tape with pre-calculated doses. […] Regardless of the patient's habitus, use the actual body weight for calculating initial resuscitation drug doses or use a body length tape with pre-calculated doses."

This guideline does recommend the use of body length tapes (without naming the Broselow tape) for emergency weight estimation. Its other message, however, is confusing and unfortunately not compatible with current evidence: TBW and IBW are not interchangeable for the purposes of drug dose calculations in underweight and obese children [61]. This guideline reflects the lack of evidence in this field.

Another way of evaluating international concepts in terms of weight estimation standards is to examine the teaching material from the resuscitation courses linked to the American Heart Association and the European Resuscitation Council. These are the major courses taught and practised in South Africa and promulgated by the Resuscitation Council of Southern Africa. The UK-based Advanced Paediatric Life Support course mentions caregiver estimates of weight, length-based methods (naming the Broselow tape and the Sandell tape) and age-based formulas, without naming a recommended system for emergency use [67]. The USA-based Pediatric Advanced Life Support course recommends the use of length-based tapes, but without naming a specific tape [68].

There are a large number of length-based tapes available in different parts of the world, most of which have never been validated in published studies. The best-known and the most widely-studied of these tapes is the Broselow tape [12]. These tapes are described in greater detail in the following sections.

In 2012, when the research for this thesis was conceived, there was already some doubt whether the Broselow tape should still be considered to be the standard for paediatric weight estimation [69]. While the initial validation studies of both the Mercy tape and the PAWPER tape referred to the Broselow tape as the current standard (or gold standard) for paediatric weight estimation, these studies found the Broselow tape to be markedly inferior in accuracy to these newer length- and habitus-based systems [70, 71]. The newer two-dimensional systems, while promising, needed

further evaluation to establish their performance in comparison to the Broselow tape which had been considered the gold-standard for the previous 20 years.

A recent systematic review on paediatric weight estimation by Young and collaborators, which was conducted independently of the Mercy and PAWPER research teams, concluded [13]:

> "Parent estimation and length-based methods with adjustment for body habitus (eg, PAWPER, Mercy) are the most accurate methods to predict children's total (actual) body weight."

At the beginning of this research, the methods setting the standard in paediatric weight estimation, based on the data available at the time, would have included parental weight estimates, the Mercy method and the PAWPER tape.

2.7 The performance of weight estimation systems

2.7.1 Overview

Weight estimation systems can be broadly divided into one-dimensional systems (based on age or length alone), two-dimensional systems (dual length- and habitus-based or dual age-and habitus-based) and other systems (parental estimations, healthcare provider guesses and miscellaneous methods).

It is fairly straightforward to see, when looking at weight-for-age or weight-for-length growth charts, that one-dimensional age- or length-based weight estimation can never achieve a high level of accuracy [11]. This is because of the large range of weights at any particular age or length. Although the range is much smaller on the weight-for-length charts, it is still substantial. From a purely mathematical perspective, the potential of one-dimensional systems is severely limited. Although the one-dimensional systems have been shown to be less accurate than the newer two-dimensional systems, they continue to be researched, used and taught [72, 73].

Two-dimensional systems, with a quantification of habitus as one of the variables, have a far higher potential for accuracy. In some studies, dual length- and habitus-based weight estimation systems have achieved nearly 90% of weight estimations

within 10% of actual weight. When assessing the accuracy of the modern weight estimation systems across multiple studies in multiple different populations, these systems manage to achieve an accuracy of 70 to 80% of weight estimations within 10% of actual weight on a consistent basis.

An important, and often ignored, confounder is that the results of weight estimation studies are very largely dependent on the composition of the study sample. Most weight estimation methods, including some age-based formulas, estimate weight very accurately in normal weight children (i.e. children close to the 50th centile of weight-for-age and weight-for-length). If a study sample has a high proportion of normal weight children, then most weight estimation systems will perform well. An example of this can be found in an article by Ralston and colleagues, in which the Broselow tape was used as comparator for another technique they were evaluating [74]. In their very large sample, 70.6% of the children were of normal weight, and the Broselow tape appeared to perform very well in the study. Only in a *post hoc* subgroup analysis was it possible to see that the tape did not perform well outside of the group of normal weight children (unpublished data). The prevalence of underweight, normal weight and overweight children in a study sample must be borne in mind when evaluating the performance of weight estimation systems.

The future development of two-dimensional or even multi-dimensional weight estimation systems will need to focus on accommodating children of all weight categories in the methodology. The current, modern systems could probably be calibrated for further accuracy.

The levels of accuracy of currently available methodology are likely to remain in the range of 80% of weight estimations within 10% of actual weight in samples with a significant proportion of non-normal weight children.

2.7.2 Age-based weight estimation

The first recorded weight-estimation system was an age-based formula developed in the early 1950s by Alexander Weech [27]. This formula was the fore-runner of the “APLS formula”, one of the most well-known of all the weight estimation formulas.

There was little progress in age-based weight estimation until the mid-2000s, when there was a surge in the number of publications on new formulas that were developed or evaluated with more than 50 publications describing or failing to validate one or more of the 20 different age-based formulas. Three of the most recent age-based systems are important to mention specifically. The “finger counting” or “Handtevy” system, is another variation of age-based weight estimation, but one which requires only recollection and not calculations (Figure 2-3). Age in years is counted starting with 1 on the thumb and counting by odd numbers to 9 on the small finger. Weight in kilograms is then counted starting with 10 kg on the thumb and counting by 5 to 30 kg on the small finger. Although this technique provides a unique way to remember weight-for-age, it is actually an analogue representation of the formula: weight = (age x 2.5) + 7.5. This method has been shown to be marginally more accurate than other formulas in two studies [75, 76]. It has failed to yield sufficiently accurate results for safe clinical use, however [77, 78].

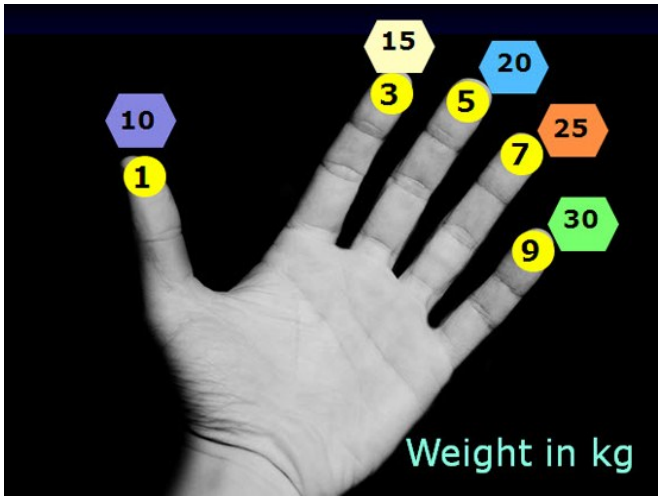


Figure 2-3 The finger-counting method. The yellow circles represent the child’s age in years. The hexagons indicate the predicted weight for each age, in kilograms. Image credit: reference [41].

The other two recent addition to age-based weight estimation are two forms of age-based, habitus-modified formulas [79]. Although the derivation studies were impressive, there has been no subsequent evaluation of these methodologies. The likelihood of achieving satisfactory accuracy with age as the primary variable is unlikely, given the inherent biological variation of length and weight for age [11]. Nonetheless, many authors continue to research age-based weight estimation in the

hope of identifying a methodology that would permit accurate weight estimation [80, 81]. One of the other reasons that the use of age-based formulas has persisted is because they are often taught on paediatric advanced life support courses [82].

There are currently more than 20 different aged-based formulas for estimating weight but to date no study has shown that age-based formulas can perform with the required degree of accuracy [78, 83]. The current evidence is not in favour of their use in clinical practice [84, 85].

2.7.3 Length-based weight estimation

The first uses of length-based weight estimation were described in 1980 and 1983, in the form of two logarithmic formulas designed to provide an estimation of ideal body weight (rather than total body weight) [86, 87]. These formulas have been shown to not estimate total body weight accurately enough, although they are more accurate than age-based formulas [88, 89]. Their presumed ability to predict ideal body weight has never been validated or studied further.

The mid-to-late 1980s saw the publication of the first tape-based weight-estimation systems as well as the first devices to provide some form of “resuscitation aid”: drug dosing and equipment sizing information for use during emergencies. The most well-known of these was the Broselow tape, but the Devised Weight Estimating Method (DWEM) and the Oakley chart were published at a similar time [8, 10, 14]. The Broselow tape is without doubt the most highly-researched of all the weight-estimation systems. It has been evaluated in both high-income countries with a high prevalence of obesity as well as low- and middle-income countries with a high prevalence of underweight children [69, 90-92]. It will be discussed in more detail in the next section. Unfortunately, the popularity of the Broselow tape, because of the resuscitation information contained on the tape, probably inhibited the further development of the DWEM system which was much more sophisticated, and accurate, in terms of weight estimation [93]. It took more than two decades for the pursuit of length- and habitus-based weight estimation to begin again. Recently both high-tech length-based weight estimation (using an electronic Wi-Fi connected

device) and low-tech table-based “look-up” systems have been explored, with accuracy results surpassing that of age-based systems [94, 95].

Several additional one-dimensional tape-based weight estimation systems have been developed, some of which are commercially available:

- The Lo rod (Hong Kong 1995)
- The Blantyre tape (Malawi 1999)
- The Pediatric Emergency Scale tape (PES tape) (Japan 1999)
- The Kloeck tape (South Africa 2000)
- The Sandell tape (UK 2004)
- The Padiatrisches Notfalllineal (PadNFL) (Germany 2004)
- The Paediatric Resuscitation Emergency Management (PREM) tape system (UK 2005)
- The Pediatape (USA approximately 2011)
- The Paulino system (Germany 2012)
- The Handtevy tape (USA 2013)

Only the Blantyre tape, the Handtevy tape, the Padiatrisches Notfalllineal, the Pediatape and the Paulino systems have been evaluated in small trials (with poor accuracy outcomes), while none of the others have published accuracy data [96-98].

2.7.4 The Broselow tape

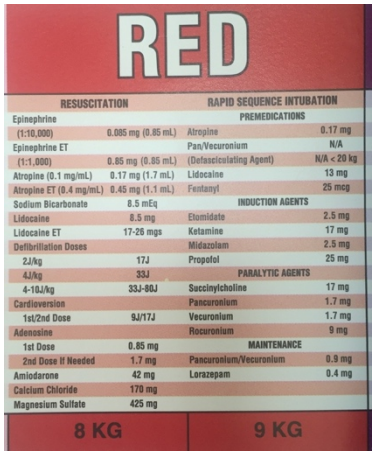
The first study on the Broselow tape was published in 1988 [10]. For many years the Broselow tape was considered to be the gold standard for emergency paediatric weight estimation. It was consistently shown to be more accurate than age-based formulas, but experts began to question its accuracy from the early 2000s [69]. Today, after more than 50 published studies on the Broselow tape, it has achieved an acceptable accuracy in only a single study from Iran, with a PW10 of 74% [99]. Many experts now consider the tape to dangerously underestimate weight in high-income populations and dangerously overestimate weight in low- and middle-income populations [69, 92].



Figure 2-4 The Broselow tape.

The Broselow can provide an estimation of weight as well as information on drug doses and equipment sizes. It is divided into nine colour zones and 26 weight categories from 3 kg to 36 kg (Figure 2-4). The drug doses and equipment sizes are provided for each colour zone and not for each weight category (Figure 2-5). The tape was first produced commercially in 1989 and has had 10 versions with the most recent being the 2011 edition A.

The tape is used to measure the supine child from head to heel – the colour-zone and weight is read directly off the tape at the point that the heel crosses the tape.


RED

RESUSCITATION		RAPID SEQUENCE INTUBATION	
		PREMEDICATIONS	
Epinephrine (1:10,000)	0.085 mg (0.85 mL)	Atropine	0.17 mg
Epinephrine ET (1:1,000)	0.85 mg (0.85 mL)	Pan/Vecuronium (Defasciculating Agent)	N/A N/A < 20 kg
Atropine (0.1 mg/mL)	0.17 mg (1.7 mL)	Lidocaine	13 mg
Atropine ET (0.4 mg/mL)	0.45 mg (1.1 mL)	Fentanyl	25 mcg
Sodium Bicarbonate	8.5 mEq	INDUCTION AGENTS	
Lidocaine	8.5 mg	Etomidate	2.5 mg
Lidocaine ET	17-26 mgs	Ketamine	17 mg
Defibrillation Doses		Midazolam	2.5 mg
2J/kg	17J	Propofol	25 mg
4J/kg	33J	PARALYTIC AGENTS	
4-10J/kg	33J-80J	Succinylcholine	17 mg
Cardioversion		Pancuronium	1.7 mg
1st/2nd Dose	9J/17J	Vecuronium	1.7 mg
Adenosine		Rocuronium	9 mg
1st Dose	0.85 mg	MAINTENANCE	
2nd Dose If Needed	1.7 mg	Pancuronium/Vecuronium	0.9 mg
Amiodarone	42 mg	Lorazepam	0.4 mg
Calcium Chloride	170 mg		
Magnesium Sulfate	425 mg		

8 KG | 9 KG

Figure 2-5 The red zone of the Broselow tape. The reverse side also contains drug information, plus equipment sizes.

There is no evidence that the drug and equipment information on the Broselow tape is beneficial without supplementary information on drug preparation, dilution and delivery, as this has never actually been directly studied. Limitations to the use of the tape include that it is only suitable for children up to 145 cm in length and that the drug dose information is incomplete and requires supplementary material for full guidance on dilution, preparation and administration.

2.7.5 Mid-arm circumference-based weight estimation

The use of mid-arm circumference (MAC) to evaluate the nutritional status (or habitus) of underweight and obese children has long been established [100-102]. In some studies, MAC has been shown to be more accurate than age or length in estimating weight, when used as a single predictive variable [54, 74, 103]. The use of MAC alone has not been able to achieve a satisfactory accuracy of weight estimation, however. The addition of MAC to multivariate models incorporating a measurement of length (or a surrogate of length) has proven highly accurate and warrants further research [74, 104, 105].

2.7.6 Guesses of weight by healthcare providers

Healthcare providers, no matter their experience, cannot guess children's weight accurately. This has been well-established to be even less accurate than age-based weight-estimation formulas [106-108]. Guesses of weight should never be used [109].

2.7.7 Estimates of weight by parents

Parental estimates of weight have been shown to be useful, and frequently accurate, as long as the parent is the regular caregiver for the child, has recently been with the child when they were weighed at home or at the clinic, and is not emotionally overwrought [107, 110-113]. Parental estimates have not consistently achieved the proposed benchmark accuracy, however, and are more likely to be incorrect in underweight or overweight children [114, 115]. A recent systematic review of weight estimation suggested that parental estimates should be one of the preferred methods for weight estimation in emergencies [13]. There is no information on how frequently parents are available at the time of initiating treatment in critically ill or injured children. The accuracy of estimations by parents in low- and middle-income countries has not been well evaluated. The limited evidence that is available suggests that parental estimates might be accurate in urban areas but estimates outside of these areas might be unreliable and not as accurate as in high-income countries [99, 116]. Alternative methods should rather be considered in these environments [9, 11].

2.7.8 Miscellaneous weight-estimation methods

There are several other somewhat more unusual methods of weight estimation that have been described previously, including formulas making use of foot size [117], or shoe size [118]. Another study evaluated the use of the age from children's clothing labels in age-based formulas, rather than their actual age [119]. This surrogate method of assessing habitus was more accurate than regular age-based weight estimation. None of these alternative systems yielded sufficiently accurate results, however.

There is one other method of weight estimation that is important. This is the "hanging-leg weight" technique described in 1990, which is significant because it

achieved one of the highest accuracy levels of any weight estimation system [120]. It was described in supine, anaesthetised children. Their lower limbs were suspended by the heels in a sling hung from a calibrated scale, with their hips flexed to 45°. The weight recorded on the scale was used in a formula to predict the actual total body weight: estimated weight = 5.176 X leg weight + 3.487. It has unfortunately not been evaluated subsequent to the initial small study. Studies such as this suggest that there might be other, more accurate weight estimation techniques that are yet to be discovered.

Since clinicians and researchers were concerned with the poor accuracy of the one-dimensional weight estimation systems, including the gold-standard Broselow tape, the possibility of estimating weight using more than one input variable was explored. This led to the development of the modern two-dimensional weight estimation systems.

2.8 Two-dimensional weight estimation systems

The two-dimensional weight estimation systems make use of two variables, most commonly recumbent length and some form of habitus assessment, to produce an estimation of weight. The DWEM was the first of the two-dimensional (or dual length- and habitus-based) weight estimation methodologies to be described [14]. It has not been updated in three decades and the data, but not the principle, can no longer be considered to be relevant. The second system to be described, the Yamamoto obesity-icon system, was first published in 2007 and is comprised of a set of length-based logarithmic regression equations for children of different age-groups and different habitus [121, 122]. A stylised icon is used to guide the user's selection of the child's appropriate habitus and the child's length is entered into the equation for that habitus category. This system has not been evaluated after the initial study, in which it performed only marginally better than the Broselow tape.

The two modern, two-dimensional systems that have shown the highest accuracy are the Mercy method and the PAWPER tape. The Mercy method has proven to be extremely accurate and has been validated in a wide variety of underweight as well as obese populations in children up to the age of 16 years [105, 123]. The PAWPER

tape has also been evaluated in a number of populations with similarly good results [70, 124-132]. In the three studies that were available in 2014, in which both the PAWPER tape and the Mercy method were evaluated, the accuracy of the PAWPER tape was slightly higher overall, although the Mercy method was more accurate in subgroups of severely obese children [130-132].

These two-dimensional systems have been shown to be substantially more accurate than the one-dimensional length- or age-based systems and will be described in more detail in the sections below.

2.8.1 The Mercy method

2.8.1.1 About the Mercy method

The Mercy method was derived from the NHANES database in the USA and makes use of a measurement of humerus length (as a surrogate for total body length) and MAC (as a surrogate for habitus) to generate a weight estimation. The humeral length and MAC are used to look up “partial weights” from a table, corresponding to the measurements. These weights are added together to obtain the estimate of total body weight [105]. There are three versions of the Mercy system - the Mercy method, the Mercy 2D tape and the Mercy 3D tape. The Mercy method requires that humeral length and MAC be measured with a regular tape measure while the 2D and 3D tapes have the partial weights printed on the tapes themselves (see below).

The Mercy method has been validated in both high-income and low- and middle-income populations and has consistently outperformed the Broselow tape [9, 71, 105, 123, 133-136]. The Mercy tapes have not been validated outside of the USA and have not performed as well as the Mercy method [71, 134]. The developers of the tape have even suggested that the system is accurate enough to outperform poorly calibrated scales [9]. The identified limitations of the Mercy method include the need for calculations to generate a weight estimate, which can be vulnerable to errors during emergencies, and its substantially reduced accuracy when used by undertrained users [71, 134]. It has also not been evaluated in the emergency setting.

2.8.1.2 How to use the Mercy method

The text below first describes the use of the Mercy method in which measurements are made with a regular measuring tape and the partial weights are obtained from a table. This is the system that is has been the most widely studied. The 2D and 3D Mercy tapes are not yet commercially available and have not been studied other than by the developers.

Step 1 – measure humeral length

Humerus length is measured in the seated child from the inferior aspect of the acromion to the tip of the olecranon. The Mercy table is used to look up the partial weight that corresponds to this length (Figure 2-6). The midpoint of the humerus is also marked.

Humeral Length (cm)	Partial Weight A (kg)	Mid-upper arm Circumference (cm)	Partial Weight B (kg)
9	0.5	10	2.8
10	0.7	11	3.8
11	0.9	12	4.6
12	1.5	13	4.9
13	2.0	14	5.3
14	2.8	15	5.9
15	3.4	16	6.5
16	4.2	17	7.4
17	5.0	18	8.0
18	6.1	19	9.4
19	7.2	20	10.9
20	8.1	21	12.4

Figure 2-6 The Mercy table (cropped) from which the partial weights can be obtained.

Step 2 – measure MAC

Mid-arm circumference is then measured at the midpoint of the arm and a partial weight obtained from the table.

Step 3 – calculate weight from partial weights

The two partial weights are added to obtain the Mercy method prediction of TBW. Limitations: The Mercy method is very accurate at estimating TBW but has no mechanism for estimating IBW in obese children. It is technically moderately challenging, and the simple arithmetic required to add the partial weights is vulnerable to errors during emergency scenarios.

2.8.1.3 The use of the 2D and 3D Mercy tapes

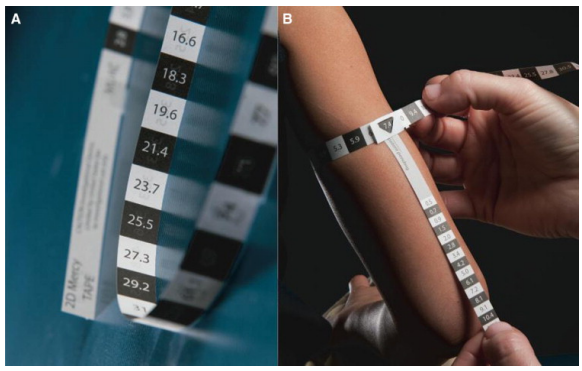


Figure 2-7 The Mercy 3D tape.

The Mercy tapes are self-contained and do not require the use of a separate measuring tape. However, neither the 2D nor the 3D versions of the Mercy tape are commercially available. The 3D tape measures humeral length and MAC simultaneously with the partial weights printed on the tape itself (Figure 2-7). The partial weights still need to be added manually to obtain the total weight estimate. The 2D tape is similar but the two measurements are made consecutively rather than simultaneously. Preliminary evidence suggests that the regular Mercy method is actually more accurate than the Mercy tapes [71].

A video on how to use the Mercy tape can be found on YouTube at: https://www.youtube.com/watch?v=9ioIp0ZOFp4

2.8.2 The PAWPER tape system

The PAWPER tape system uses actual total body length and a gestalt impression of habitus to generate the weight estimate. A subjective visual assessment of habitus is used to classify the child into a habitus score category. The weights for each habitus score can then be read directly off the tape at the segment corresponding to the child's length [70].

2.8.2.1 The development and science of the PAWPER tape

There was a two-part incentive to create the PAWPER tape: firstly, the existing weight-estimation systems at the time were just not accurate enough (including the "gold standard", the Broselow tape); and secondly, the Broselow tape was very expensive. There were many studies that showed weight estimation by the Broselow tape to be dangerously inaccurate: specifically, in children in the upper centiles of weight-for-length (obese children) and now, more recently, in children who are in the

lower centiles of the weight-for-length charts (underweight children). The PAWPER tape was devised to optimise the accuracy of weight prediction in children through including a formalised estimation of body habitus into the standard length-based weight estimation. The PAWPER tape (Figure 2-8) and subsequent systems have always been produced on a not-for-profit basis. The tapes are sold only to cover the cost of production.

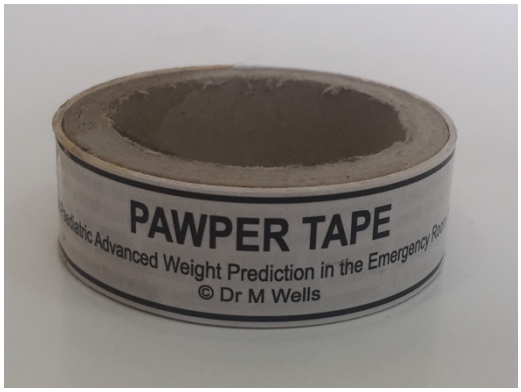


Figure 2-8 The PAWPER tape. The tape is wound on a sturdy cardboard reel for protection.

2.8.2.3 The science behind the PAWPER tape

The PAWPER tape was developed at the University of the Witwatersrand in 2007 and validated in a preliminary study in South Africa [70]. This system makes use of length-based, habitus-modified weight estimation. A weight predicted by length can be modified up or down based on a visual assessment of the child's body habitus: the weight of overweight children is adjusted up and the weight of underweight children is adjusted down. The name is a word play on 'pauper' and is a reference to the very high cost of the Broselow tape (especially in low- and middle-income countries) compared to the economical and affordable price of the PAWPER tape.

The PAWPER tape is, in essence, a full-scale approximation of the 50th centile of the World Health Organisation (WHO) weight-for-length growth chart. Each weight division corresponds to the 50th centile of length for that weight, with this centile falling in the centre of each weight division on the tape.

Since not every child has a weight on the 50th centile for their length, the PAWPER tape methodology included a mechanism to improve the accuracy of the weight estimation of children with a higher or lower than average weight-for-length. Since it had been shown previously that weight can be predicted more accurately when a measure of frame-size was used to modify the length-based prediction methodology,

a tape-based system was devised whereby the weight predicted by length could be modified by using a five-point assessment of body habitus [14]. The body habitus of each child, as appropriate for their age, is evaluated by rapid inspection and then categorised as a habitus score:

- Habitus Score 1: very small frame and/or very thin with ribs easily visible.
- Habitus Score 2: slightly smaller than average frame and/or decreased adiposity ("thin" / "petite").
- Habitus Score 3: average frame size and optimum body habitus.
- Habitus Score 4: slightly larger than average frame size and/or somewhat increased adiposity ("chubby" / "well padded").
- Habitus Score 5: very large frame and/or significantly increased adiposity, tending towards obese ("fat" / "overweight").

This appraisal of body habitus was deliberately based on general observation or gestalt impression rather than any specific anthropometric or morphological assessments. This was because it needed to be a process that could be rapidly employed and easy to perform in order to avoid delays in weight estimation and, consequently, emergency medical treatment.

Once an estimation of body habitus and a body habitus score has been established, the modified weight-for-length can be read off the PAWPER tape. Weights approximating the 5th, 15th, 85th and 95th weight-for-length centiles at that specific length are shown in parenthesis for each weight division (Figure 2-9). These weights correspond to the predicted weights for each habitus score (i.e. 1, 2, 4 and 5) at that length (see below – how to use the PAWPER tape).



Figure 2-9 A sample segment of the PAWPER tape. This shows the different weights for the five Habitus Scores. The Habitus Score 3 weight is shown in the large font.

2.8.2.4 How does the PAWPER tape differ from the Broselow tape?

In many ways it is easiest to describe the features of the PAWPER tape by differentiating them from those of the Broselow tape:

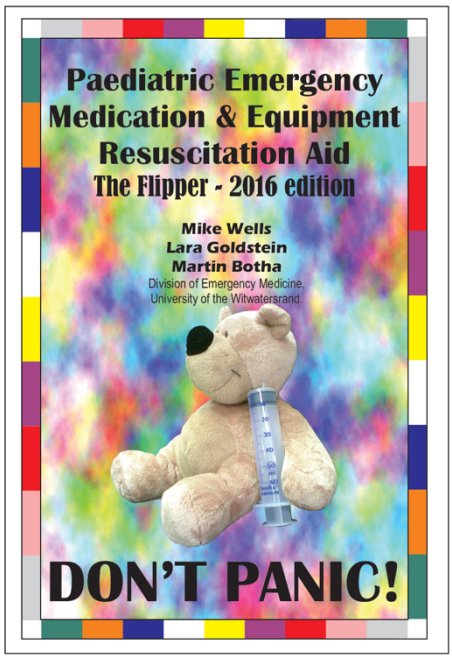


Figure 2-10 The Flipper is a low-cost drug dosing guide for children.

1. In the patent applications, the Broselow tape is described as primarily a device to predict medication doses and equipment size based on a child's length. The PAWPER tape is designed solely to estimate weight in children with a recommended companion book for drug dosing information, "*Emergency Drug Dosing in Children*" [1], or the dosing pocket-card, "The Flipper" (Figure 2-10).
2. The Broselow tape is a multi-use device, and costs about R500. The PAWPER tape is a single use, disposable device, available at a cost of less than 1 US dollar (R5).
3. The Broselow tape is based on the American CDC NHANES 2000 growth charts. The PAWPER tape is based on the WHO growth charts.
4. The PAWPER tape reflects the 50th centile in the centre of each weight division; the Broselow tape reflects the 50th centile at the distal edge of the weight division. In other words, the PAWPER tape would predict an actual weight of 13.8 kg to be 14 kg, while the Broselow tape would be more likely to predict an actual weight of 13.8 kg to be 13 kg. The PAWPER tape also provides higher estimations of weight than the Broselow tape at lower measurements of recumbent length (Figure 2-11).
5. The Broselow tape has a very limited, untested mechanism for increasing the weight estimation accuracy in children who are obese, but none for underweight or small for length children [137]. The PAWPER tape included approximations of weights at the 5th, 15th, 85th and 95th WHO weight-for-length centiles, with a method to adjust the weight according to the frame size or degree of adiposity of the patient.

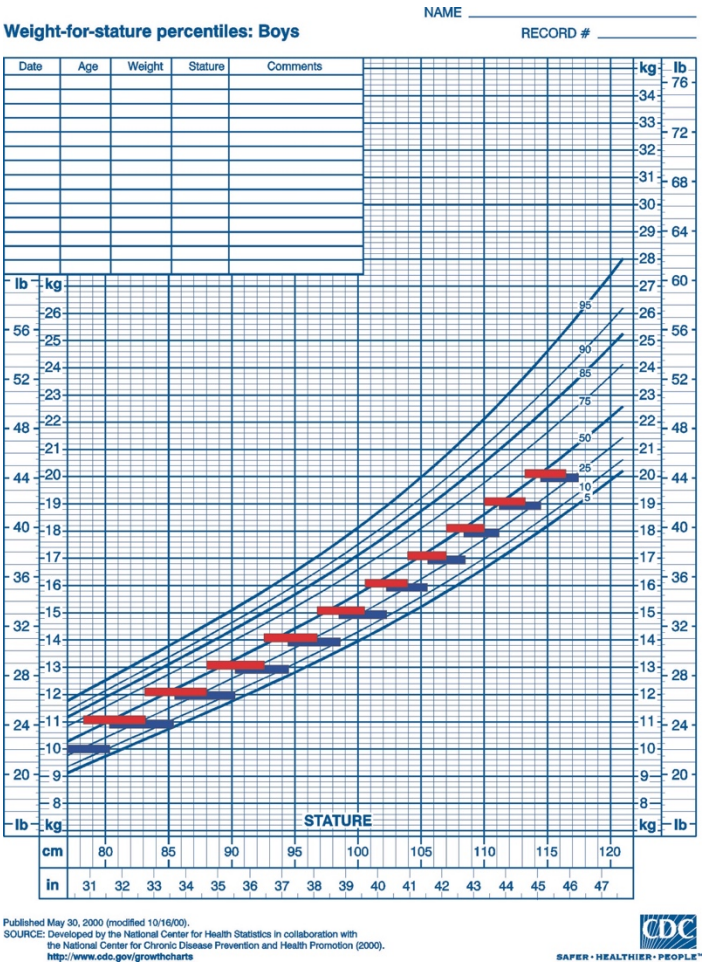


Figure 2-11 The relative positions of the tape weight divisions for the Broselow tape (blue) and PAWPER tape (red).

The PAWPER tape makes use of the same weight divisions found in the Broselow tape in order to maintain standardisation in the use of dosing guides and adjuvant materials.

2.8.2.6 How to use the PAWPER tape

A video describing the use of the PAWPER tape may be found on YouTube: https://www.youtube.com/watch?v=FenVSBzyBv8&t=8s

Step 1 – measure length

When using the PAWPER tape, measure from the child's head to his or her heel, starting with aligning the end of the tape that reads "MEASURE FROM THIS LINE" with the top of the child's head (Figure 2-12). Stretch the tape along the length of the

child's body in a straight line and read off the weight at the point at which the tape crosses the heel (Figures 2-13 to 2-15).

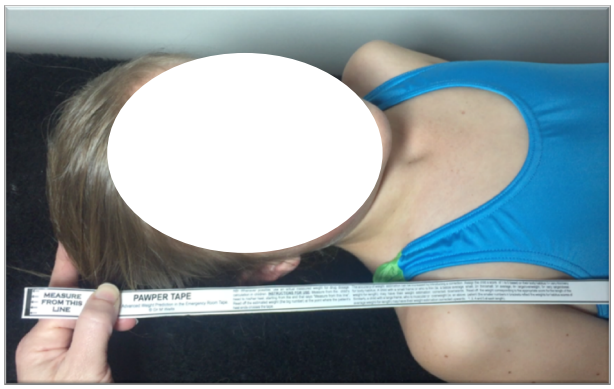


Figure 2-12 Position of the tape at the child's head. The beginning of the tape is aligned with the top of the child's head.

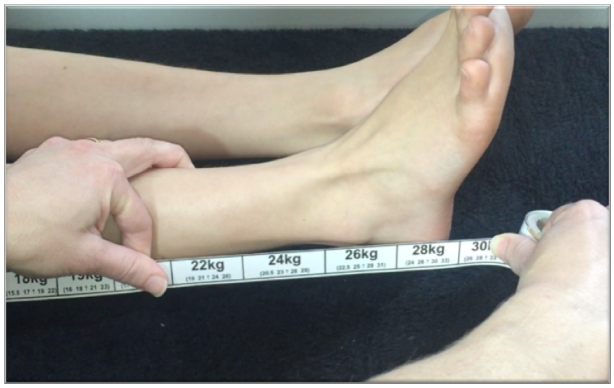


Figure 2-13 Position of the tape at the child's heel. The point at which the tape crosses the heel indicates the length-segment into which the child falls.

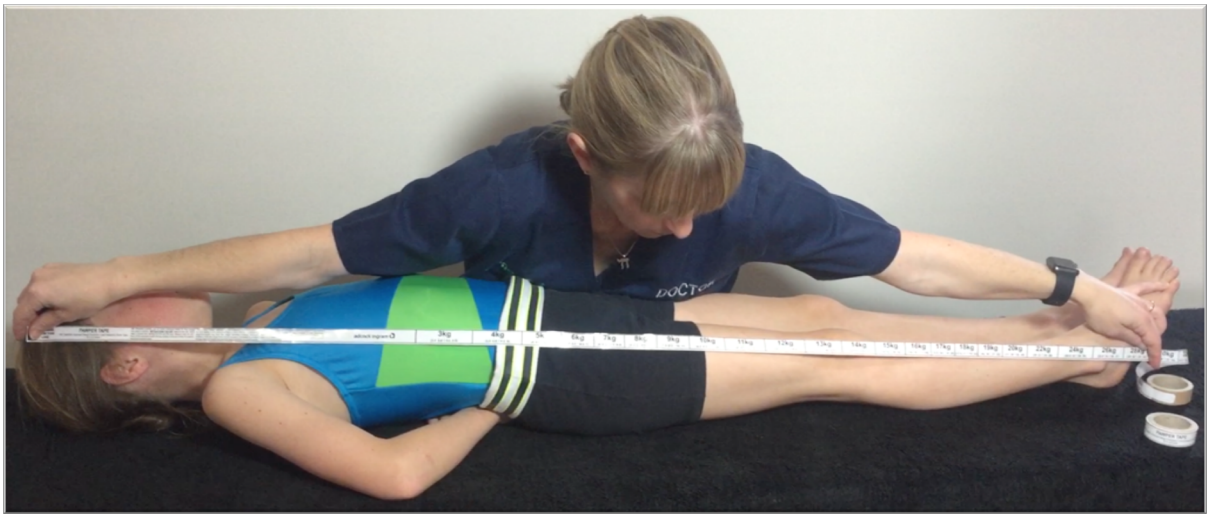


Figure 2-14 The correct measurement technique for the PAWPER tape – side view. The tape should be kept straight and must not follow the contours of the body. The tape should be on the same side as the user – this image is merely for demonstration purposes.

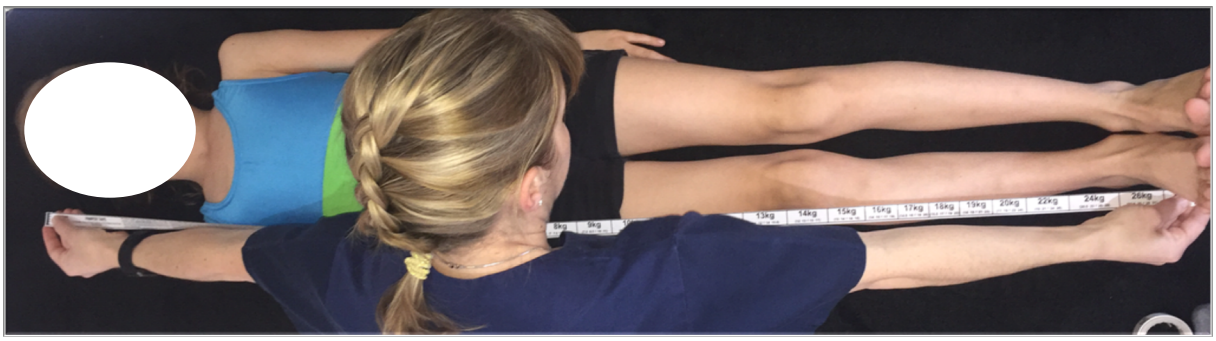


Figure 2-15 The correct measurement technique for the PAWPER tape – top view. The user must ensure that the tape is correctly aligned at the head and the heel for an accurate measurement.

Step 2 – identify relevant length segment

This weight category determined by the child's head-to-heel length may be used for drug dose calculations in a normal weight child of "average" build (Habitus Score 3) (Figure 2-16).

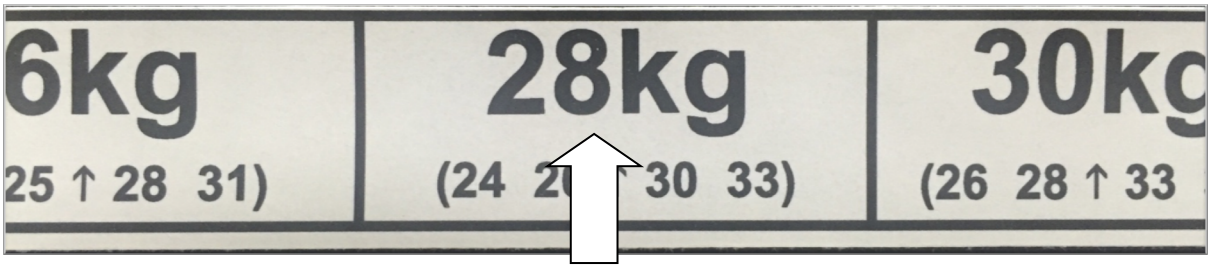


Figure 2-16 A length segment of the PAWPER tape showing the HS3 weight. The white arrow indicates the weight of a child with "average" habitus. This is indicated on the tape with a large font.

Step 3 – assign habitus score

If, however, you think that the child is heavier or lighter than average for their length, the weight estimate can be adjusted by the use of a habitus score. Assign the child a score from 1 to 5 based on the size of their frame and their degree of adiposity.

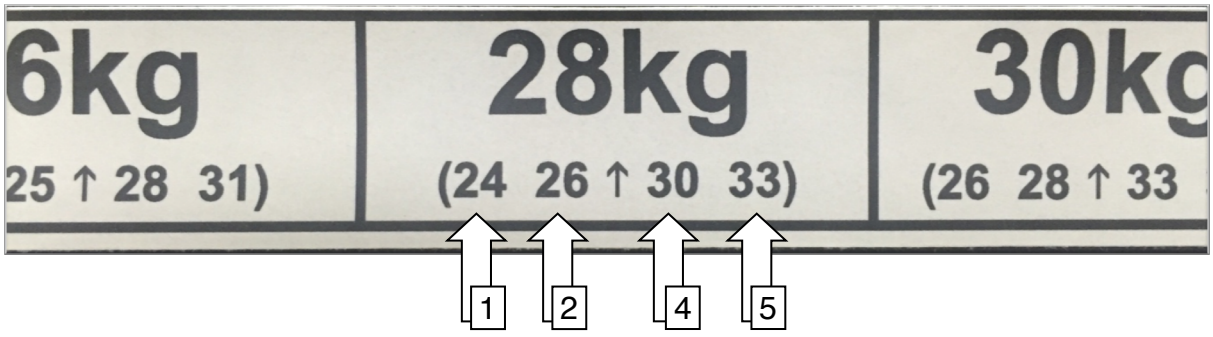


Figure 2-17 A length segment of the PAWPER tape showing the other HS weights. The white arrows with numbers point to the predicted weight for children with a habitus score of 1 and 2 (below "average" or "normal" weight) as well as 4 and 5 (above "average" weight).

The numbers in brackets on the PAWPER tape correspond to the weight estimations for each habitus score at that length (Figure 2-17). The weight can be estimated with a much higher degree of accuracy by using the habitus-modified weight in children with a higher or lower than average weight-for-length. A child with a Habitus Score of 3 (average weight or normal weight) has an unchanged weight estimation (use the large-font number).

Step 4 – drug dosing companion

Use a reference source, such as a computer-based or chart-based system, to look up the equipment size, drug dose and drug preparation information (Figure 2-18). The reference should comply with evidence-based recommendations on the presentation of drug dose information to minimise medication errors in children - the drug dose information should include the dose based on body weight, the total drug dose, the method of dilution and the millilitre volume of the diluted drug to administer to the child.

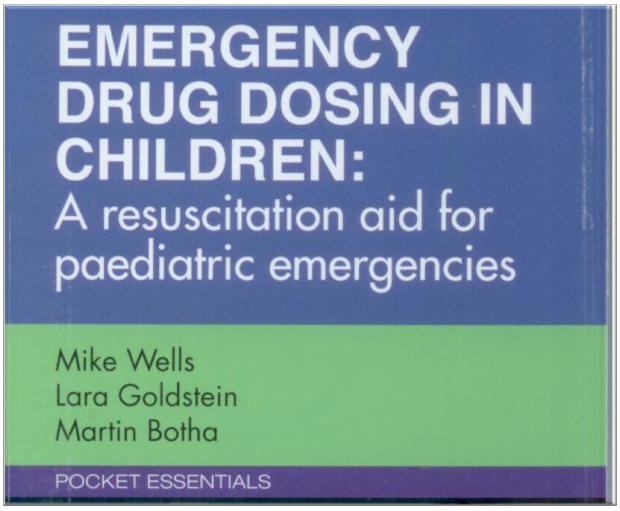


Figure 2-18 A cropped image of the cover of the Emergency Drug Dosing in Children book cover.

2.8.2.7 Current use of the PAWPER tape system

Is the PAWPER tape being used in clinical practice?

The PAWPER tape is in wide-spread use in South Africa in Emergency Departments in both the public and private sector. It has, in fact, been recommended by the Ideal Clinics program of the National Department of Health as an alternative to the Broselow tape and is part of the paediatric section of the South African National Treatment Guidelines [138, 139]. The PAWPER tape is also used by many public and private prehospital advanced life support provider companies.

Has the PAWPER tape been endorsed?

The PAWPER tape has been supported by the Emergency Medicine Society of South Africa (EMSSA), the Paediatric Emergency Care special interest group of the Emergency Medicine Society of South Africa and the Expert Consensus Report for Emergency Centres in Western Cape [140]. It is also endorsed by the African Federation for Emergency Medicine and the International Federation for Emergency Medicine for weight estimation in children [141]. The PAWPER tape is used and

taught on the Advanced Paediatric Life Support (APLS) and Pediatric Advanced Life Support (PALS) courses in South Africa.

Have the PAWPER tapes and EDD4C app been validated?

Most of the prospective studies on the PAWPER tape systems have been performed in Johannesburg, South Africa. There have been no studies on the tapes outside of Johannesburg within South Africa, but it has been extensively studied and independently validated internationally. Studies have been, or are being, performed in the USA, Australia, Nepal, Turkey, Belgium and Rwanda [142]. In addition, two of the "virtual" studies have been performed using international datasets - one from the NHANES dataset in USA and one from a dataset of data pooled from multiple low- and middle-income country nutritional surveys [74]. The mobile phone app was developed and evaluated as part of this research. It needs further development and refinement before it can be distributed for widespread clinical use. It also faces technical regulatory issues with regards to receiving approval to be made available on the Apple platform.

Physical properties of the PAWPER tapes

The original PAWPER tape is made of a non-stretchable paper tape. Some departments have mounted it between acrylic glass (Perspex) sheets to eliminate the potential for damage and to allow cleaning. It is a very low-cost item (about 50 US cents) and should ideally only be used once as the paper tape cannot be cleaned. The first version of the PAWPER XL tape was made from an encapsulated, heavy-weight paper tape to allow for repeat usage (the same material as the Broselow tape). The current version of the PAWPER XL tape is printed on thin plastic material that is non-stretchable. It can be used multiple times and the manufacturer has certified that the material will not stretch with normal repeated use (although this has not been formally tested as yet). The PAWPER XL tape can be wiped clean as it is made of plastic but should be discarded if it becomes contaminated with body fluids. This plastic material is more easily cleanable than the plastic-encapsulated paper version.

2.8.2.8 Limitations of the PAWPER tape

Several studies provided some understanding of how and why the PAWPER tape performed less well than expected in some populations and also provided some insight to how it may be improved [127, 128, 130]. The three important limitations of the original tape were that it was not long enough for all children, that the assessment of habitus was too subjective and that it did not make adequate provision for obese and severely obese children [127-129, 143]. These factors are discussed below.

Too many children "too tall for the tape"

The original PAWPER tape can only accommodate children up to 153 cm in length which may result in an inability to provide a weight estimation for taller children [143]. This limitation is one of the major criticisms of the Broselow tape [143]. In some studies, up to 30% of children enrolled had to be excluded for exceeding the length of the Broselow tape [88, 130, 143, 144]. Although the PAWPER tape is slightly longer than the Broselow tape (153 cm vs 145 cm), this limitation remains significant.

Since the Mercy methodology has established the potential to provide an accurate weight estimate in children up to 16 years of age, this has set a standard that should be emulated by other techniques [105]. Based on an analysis of the NHANES growth charts, increasing the PAWPER tape length to 180 cm would accommodate more than 95% of children up to the age of 18 [145].

Too many children "too fat for the tape"

The results of two studies in obese USA populations showed an underestimation of weight in obese children [127, 128]. While this was partly due to errors in habitus assessment, the PAWPER tape was also not able to provide weight estimates for severely obese children. Since obese children are potentially more at risk of poor outcomes following drug dose errors than normal weight children, this limitation is significant (see section 2.11) [49].

Misclassification of habitus score

The PAWPER methodology makes use of a gestalt assessment of habitus to allow for rapid estimates of weight [70]. Although this system is simple, it is vulnerable to biases in assessing habitus, especially in obese children [146]. It is also susceptible to differing interpretation of habitus based on the prevalence of overweight and obese children in the population [129, 147].

The assessment of habitus is reviewed further in the next two sections.

2.9 What is the best way of assessing children's body habitus?

The accuracy of visual, gestalt, or anthroposcopic assessment of body habitus has been established in the somatotyping as well as the weight-estimation literature, but it remains subjective and at least somewhat dependent on the experience of the assessor [95, 148-157]. Two previous studies showed inaccurate habitus assessment to be the main contributing cause of weight estimation error by the PAWPER tape [128, 129]. The original inspiration for the concept of the habitus score was derived from the techniques of body condition scoring found in the veterinary literature, many of which make use of reference images for training and to guide assessment [158-176]. In humans, the benefit of figural reference images to assist with habitus assessment has been proven in body image and nutritional research [177-182]. These figural reference images, also referred to as silhouette drawings or figural stimuli, consist of a series of seven to nine frontal images of a range of thin to fat children [178]. These images can be used by an assessor to match actual habitus to a reference image. This concept is more cognitively suited to emergency use than checklists or visual scales [183]. A preliminary study with the PAWPER tape using this system (a set of modified Collins figures, Figure 2-19), showed the use of reference images to be as accurate as gestalt assessment of habitus by an experienced user [131, 179].

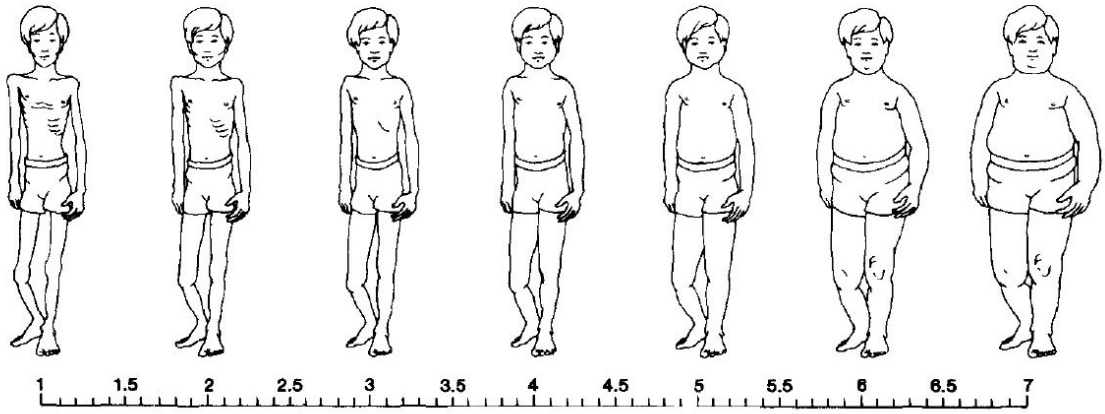


Figure 2-19 The Collins figural images for boys.

This could potentially allow a greater standardisation and objectivity of habitus assessment but needs further verification.

2.10 Can anthropometry be used to assess habitus?

Although figural reference images might improve the accuracy of habitus assessment when compared to unaided visual estimation, there is still an element of subjectivity to the assessment [182]. The use of a fully objective method of assessing habitus, such as anthropometry, could eliminate this problem. There is evidence that some simple anthropometric measurements, most notably MAC, are strongly associated with habitus [54, 103]. Not only has MAC been shown to be predictive of habitus and total body weight in children, but it has been shown to be useful in weight estimation as part of the Mercy method [105]. Measurements of MAC have been shown to have a high inter-rater agreement, an essential requirement for the habitus assessment to be consistently accurate [184].

Although the use of anthropometric measurements to classify habitus may be more time-consuming than a gestalt or image-assisted assessment of habitus, it could improve the standardisation of weight estimations by the PAWPER tape across different populations. The use of anthropometric techniques, however, would have to be in a manner that would not require any form of calculation to mitigate cognitive overload.

2.11 How do cognitive demands affect weight estimation?

An understanding of the cognitive impact of weight estimation systems is relevant because it determines their usability, their vulnerability to error (and therefore accuracy) and their suitability for use during emergency care. The cognitive load experienced by healthcare providers during resuscitative care decreases their mental acuity and increases the risk of incurring errors of all types [25, 36]. While the healthcare provider's major mental focus should be on the details of the medical management of the child, there are additional unavoidable cognitive stressors associated with treating children. These primarily involve the process of weight estimation and drug dose calculation and preparation for delivery [1]. The negative effects of this additional and distracting "thinking time" must not be underestimated, nor should the ability of healthcare providers to perform even simple tasks without error be taken for granted [185-187]. The process of weight estimation during the management of medical emergencies in children is, therefore, not straightforward and the methods to evaluate and improve weight estimation techniques should take this into account.

The factors that ultimately determine the accuracy of weight estimation in the emergency medical care of children are as follows:

- Inaccuracies in the calculations required in the use of a weight estimation system may result in errors. These errors are more likely to occur during emergency situations [188-190]. Even simple arithmetic calculations are prone to error during emergencies. Any system, therefore, that requires calculations is vulnerable to mistakes during emergencies, even if a calculator is available [190-196].
- Additional "thinking" tasks (such as recalling formulas from memory) may impair the mental work of managing acute medical care, resulting in cognitive overload, with consequent risk of errors in medical management [25, 197-200]. Memory is capricious and unreliable during the stress of emergencies, irrespective of the experience of the healthcare provider. Over-reliance on memory and the under-use of technology (in various forms) may increase the

risk of errors as well as the cognitive load [201-203]. Thus, any system that requires the recollection of formulas or doses is susceptible to errors and will generate significant cognitive stress [204-206]. The use of technology can decrease error and cognitive burden related to all aspects of drug dosing [207-210].

- In order for any system to function effectively, users must be sufficiently trained, ideally with simulation training under conditions as similar as possible to real-life. Simple education without practice is inadequate [134, 211-213]. Unless sufficient instruction and practice is attained in the use of weight-estimation systems in simulation training scenarios, users may experience substantial errors [185-187, 214-216]. Even the simplest of methods to use, such as the Broselow tape, may be used incorrectly by the poorly-trained user [185].

Since the ultimate goal of weight estimation is to enable safe and effective drug dose reckoning, it is imperative that a weight estimation system be designed not only to minimise cognitive burden, but also to integrate effectively with a drug-dosing resuscitation aid [98]. An appropriate balance between accuracy, time-to-perform and cognitive demands is required [1]. While it is clear that weight-estimation methods requiring any form of recollection or calculations are unreliable, length-based weight estimation by means of a tape system has been suggested to be relatively neutral in terms of cognitive burden [25, 32]. Furthermore, although there is no direct evidence on the impact of assessing habitus on cognitive burden, the type of cognitive load generated by a rapid gestalt assessment of body habitus or the use of figural reference images has been suggested to be low [183, 206].

2.12 Should weight estimation be modified in obese children?

Obesity in childhood is problematic in many countries across the world, including South Africa [217]. The “epidemic of obesity” remains a concern because, although the prevalence of obesity has plateaued since 2009, the prevalence of *severe* obesity has continued to rise [218]. In low- and middle income, obesity has become part of the “double burden of malnutrition”, with high prevalences of obese and underweight children in the same population [219]. Even in Africa, the obesity rate for children

under 5 years of age is significant, ranging from 6.5% in the Southern region to 26.6% in the Northern region, averaging 12.7% across Africa [220]. In children aged 5 to 19 years, the lowest prevalence of overweight and obese children is in Burkina Faso at 7.6%, with the highest in Algeria at 31%, while only seven countries in Africa have a prevalence of less than 10%. Southern Sudan, described as the “hungriest place on earth” paradoxically has a childhood obesity rate of 12.6% [92, 221]. Furthermore, the rise in childhood overweight and obese children in low- and middle-income countries is increasing at a rate 30% faster than in high-income countries [221].

The prevalence of childhood obesity is high enough in many populations, including those from low- and middle-income countries, that it is advisable for weight estimation protocols to include a mechanism to deal with obese children, as they are at increased risk for medical errors [49, 222-224].

It is only in the last several years that some degree of consensus has emerged on the appropriate dose-scaling of emergency drugs in obese children i.e. which weight-descriptor to use for which drugs. In general, bolus doses of hydrophilic drugs should be dosed to IBW in obese children and initial bolus doses of lipophilic drugs should be dosed to TBW, with some exceptions [5, 64, 225, 226]. This means that, in order to avoid medication errors in obese children, the ability to estimate both TBW and IBW is essential to ensure patient safety [55]. This means that drug dose determinations are significantly more complex in obese children than in non-obese children, but it is essential for these doses to be precise, as incorrect dosing in obese children has been suggested to be a significant cause of excess morbidity and mortality when compared to normal weight children [224, 227-230].

Some experts have argued that IBW may be used for dose-scaling for all children as this would avoid overdosing obese children with hydrophilic drugs [53, 55]. This is now known to be incorrect, however, as obese children would then be underdosed with lipophilic medications if IBW was used as scalar instead of TBW [231]. It is also clear that IBW is not an appropriate dose-scalar for drug dosing in children who are not obese, as this may result in dangerous overdosing in underweight children [1, 130, 231]. While the need for variable dose scaling in emergency care is evident, the

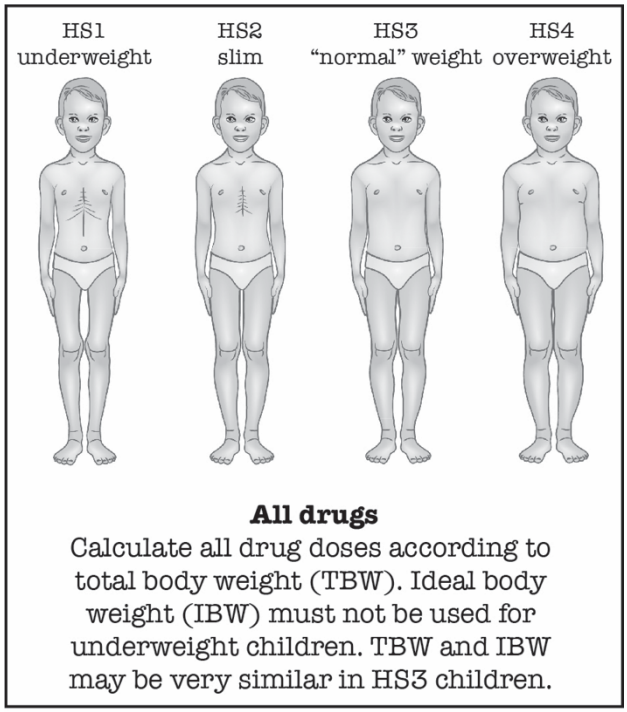
impact on patient outcome and on the complexity of the drug dosing processes has not yet been fully elucidated. There is currently very little published data on the use or estimation of IBW during emergency care. The length-based formulas, which were originally intended to estimate IBW (the Traub-Kichen and Traub-Johnson formulas), may prove to have value in the future for this purpose [86, 87].

The American Heart Association guidelines for paediatric advanced life support reflect the uncertainty of the available evidence:

> "Despite these theoretical considerations, there are no data regarding the safety or efficacy of adjusting the doses of resuscitation medications in obese patients. Therefore, regardless of the patient's habitus, use the actual body weight for calculating initial resuscitation drug doses or use a body length tape with pre-calculated doses" [66].

Their guideline essentially advocates for the use of either TBW *or* IBW (from length-based tapes). If both weights are easily obtainable, it seems reasonable to dose drugs according to the existing pharmacological consensus until further evidence emerges.

A list of hydrophilic and hydrophobic (or lipophilic) drugs that are often used in the Emergency Department are shown in Figure 2-20. This figure illustrates the appropriate weight descriptor for obese and non-obese children and the applicable weight-descriptor to be used (adapted with permission) [61]. IBW should only be used as a weight descriptor for obese children, to scale the doses of hydrophilic drugs. There may be some differences between disciplines in the preference for dose calculation of some drugs, e.g. non-depolarising muscles relaxants might be dosed to IBW by anaesthetists to ensure a controlled duration of action, while emergency physicians may prefer dosing to TBW to ensure quick onset and long duration of action.



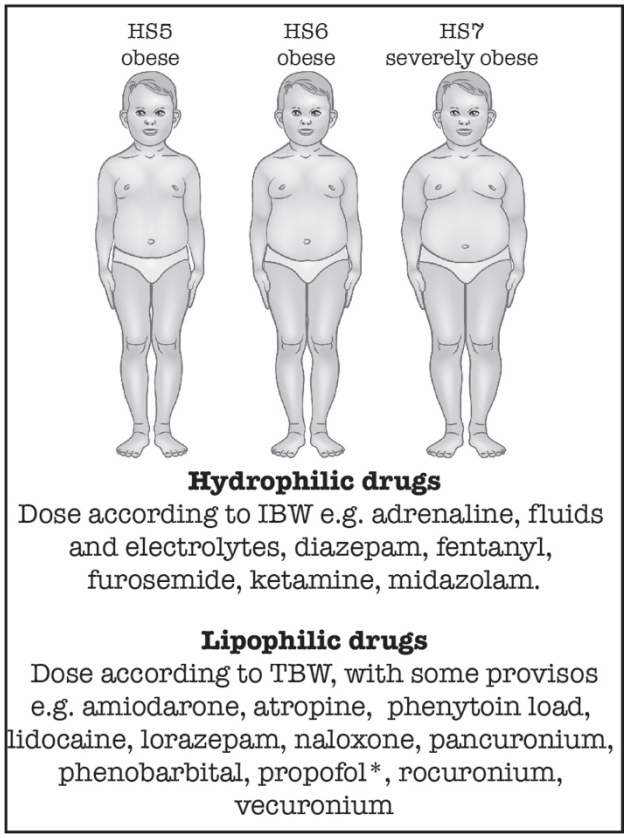


Figure 2-20 Weight descriptors to use for underweight, normal weight and obese children. *Propofol should, theoretically, be dosed to TBW but many obese children exhibit an increased sensitivity to the drug – reduced or titrated doses should be considered.

2.13 Can mobile phone applications improve weight estimation?

While the future must bring a major swing towards the use of technology during the resuscitation of children, there will always be parts of the world that will not have access to this technology. Both high and low technology systems will always be required to ensure that all children receive optimum care [1].

There is very limited research on the value of mobile phone apps during paediatric emergency care, but the little information that does exist supports the value of well-designed smartphone apps [207]. There are, however, currently no apps on the Apple iStore or the Google Play store that provide for weight estimation in children during emergencies.

There is no doubt that the use of apps could be of tremendous value in paediatric resuscitative care. The “black box” model (Figure 2-21) would allow the use of complex mathematical calculations, at virtually zero cognitive cost. The key to the black box model is that the user needs only to input very simple information and the app could provide information on weight, weight status and drug dosing information with no additional work or input from the user [232]. The "black box" portion of the system contains rules, formulas and calculations that the user does not see or need to know in order to use the system. For example, the user may enter the child’s length and habitus score. The app will calculate the TBW and IBW and will automatically calculate drug doses, method of preparation and correct volume to administer based on whether the child is underweight, average weight, overweight or obese. This would result in a dramatic reduction in thinking time and potential for errors [207].

The other potential uses for smartphone apps are the automatic measurement of a child’s length, without having to measure it, and the automatic classification of habitus. Preliminary work on the automatic measurement of length and its implications for weight estimation has been promising but will require further validation [233].

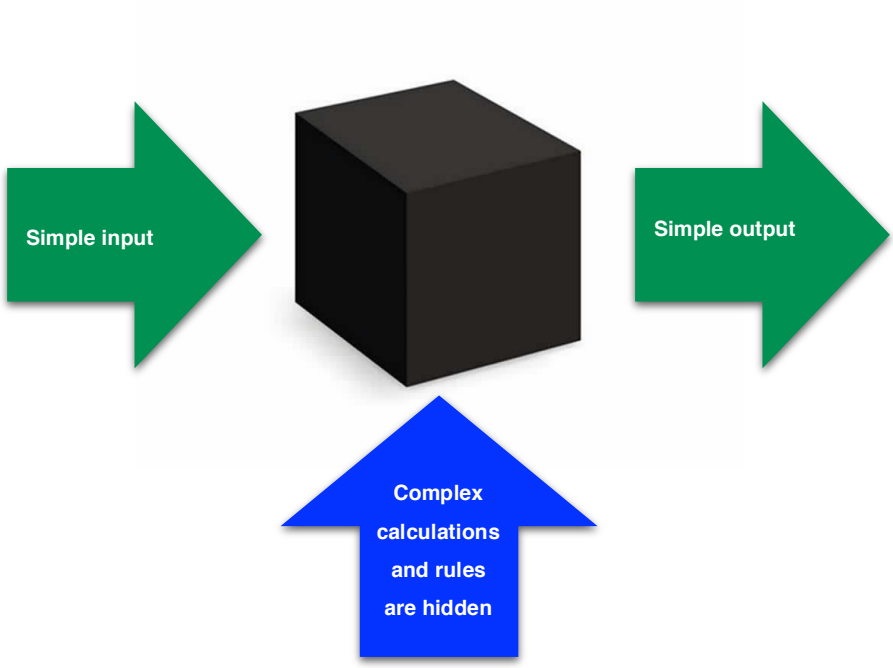


Figure 2-21 The black box model adapted to cognitive aids.

The accuracy of a smartphone app to digitally analyse whole-body silhouettes to predict body composition in humans has recently been verified [234]. This may potentially be able to be used to estimate both TBW as well as body composition in an emergency setting in the future. The use of automated systems to assess condition score (habitus) in animals has already been well-established and developed on a commercial scale [170, 176]. This technique is, therefore, potentially feasible in humans.

2.14 How do real-world factors cause weight estimation errors?

There are three contributors to inaccuracies in emergency weight estimation: system errors, human factor errors and patient factor errors.

System errors derive from the inherent performance of the weight estimation system. The performance of the system under ideal circumstances will depend on the population in which it is used and its inherent accuracy [13]. This aspect of weight estimation is the most researched and the best understood (see section 2.7). Human factor errors are defined as errors resulting from poor or incorrect use or application of the weight estimation system [185]. Even the simplest weight

estimation systems, that are straightforward to use such as the Broselow tape, are prone to error and incorrect usage if the user has not received sufficient training and practice with the system [185-187, 235, 236]. Systems that require the recollection of a formula and the accompanying arithmetic are also hazardous and error-ridden [195, 205]. Weight estimation errors as high as 20% to 50% have been ascribed to human factor errors [134]. The evidence available in this field, although still scanty, has generated important concerns about patient safety during emergency care, if the treating team is not appropriately trained and prepared for appropriate weight estimation [134, 212].

With respect to patient factor errors, no research has been conducted on the effects of different patient positions and patient cooperation on the accuracy of weight estimation. While some authors have acknowledged the potential importance of these factors, their significance is unknown [237]. It is absolutely certain that weight estimation is required in many children who are not supine and not cooperative. It is extremely unlikely that weight estimation systems will be as accurate in these children as they may be under ideal circumstances. This is an important consideration that demands further research.

Both human and patient factor errors may be resolved with sufficient training, however, there is no standard for instruction in weight estimation. None of the major paediatric emergency care training organisations have guidelines on the incorporation of weight estimation and drug dosing training into resuscitation simulations [238, 239]. This is unfortunate as the weight estimation process may ultimately significantly influence patient safety. It is likely that the process of weight estimation, drug dose calculations, drug preparation and simulated administration is seldom practised to the level that would be required in a real resuscitation. This also requires further research.

2.15 Summary

There are very large knowledge gaps in this field. While much is known about which weight-estimation methods are not sufficiently accurate, not enough is known about the newer two-dimensional systems and how their accuracy can be improved.

Virtually nothing is known about how these systems actually work during emergency care as well as about the consequences of inaccurate weight estimation. In order to be faithful to the imperative to “do no harm” to our patients, these matters need further research.

2.16 Gaps in the literature

At the beginning of this research process, around 2012, there were many significant gaps in the literature pertaining to paediatric weight estimation and relevant related topics. Some of the important area of knowledge deficits are shown in Table 2-2 below.

Table 2-2 A summary of the gaps in the paediatric emergency weight estimation literature.

Major relevant knowledge gaps at the beginning of this research	Existing evidence
Accuracy of weight estimation systems	
Which weight estimation systems are the most accurate?	Although there have been multiple studies evaluating or comparing the accuracy of methods of paediatric weight estimation, there have been no systematic reviews or meta-analyses on the topic*. There has also been no clear consensus on this question. *One systematic review was published in 2016, after the beginning of this research.
Are parental estimates of weight valid in all populations?	There have been several studies on the accuracy of parental estimates of weight in urban areas of high-income countries, there has been no research on this in rural areas and low-income countries. There is substantial reason to believe that this method of weight estimation might not be valid outside of high-resource areas.
Do weight estimation systems perform differently in high-income and low- and middle-income country populations?	Multiple studies have been conducted in both high-income and low- and middle-income country populations, but little work has been done on the integration of these findings. There has been no published work dedicated to systematically evaluating the role of weight estimation systems in these different populations.

Does the accuracy of weight estimation change outcome: is an accurate estimation of weight really necessary?	This issue has been frequently raised in publications but has never been addressed in appropriate prospective research.
How accurate should weight estimation systems be?	There have been limited comments within publications speculating on the appropriate accuracy for weight estimation systems, but this has never been addressed directly in the form of reviews, surveys or prospective research. Only one study (masters research report) published in 2009 suggested a target accuracy for weight estimation systems.
Can assessment of habitus improve weight estimation?	There has been minimal research in this field – only the DWEM system (the devised weight estimating method) has been researched, other than the preliminary work on the PAWPER tape. The DWEM makes use of a visual assessment of habitus but has been evaluated in only two previous studies. *Multiple studies have since been conducted after this research was commenced.
Could the PAWPER tape be expanded for better performance in taller and obese children while maintaining accuracy in underweight and normal weight children?	The PAWPER tape has not been fully evaluated. It is restricted to 153 cm in length. There is no published evidence about an extended capability in taller children or obese children.
What effects the accuracy of visual estimation of habitus?	There has been no formal evaluation of the accuracy of visual assessment of habitus, the factors that might affect this accuracy or the training that might be required to assess habitus accurately. There is limited data available on the use of reference images to assist in habitus assessment in the body image analysis field.
Which method of assessing habitus is most accurate?	There has been no previous work comparing different methods of assessing habitus for the purposes of weight estimation.
How many habitus score categories are required for accurate weight estimation?	There has been no previous work to evaluate the number of habitus categories to allow for accurate weight estimation. The work done with animals in the veterinary

	field and the work done in body image analysis has provided some evidence, but this is based more on speculation than research-based evidence. *Some work has been published in the last two years on this topic.
Do different categories of healthcare providers have differing ability to assess habitus?	This has never been studied before and there is no evidence to support this suggestion. There is only anecdotal evidence from several sources, to suggest that visual assessment of habitus, for the purpose of somatotyping, improves with training and experience.
Can anthropometric techniques be used to aid weight estimation?	There is some evidence to show that mid-arm circumference can be used to estimate weight, albeit with a low degree of accuracy. There is no evidence available that combining anthropometric measurements with other techniques can aid weight estimation (other than the Mercy method).
Can technology be used to aid weight estimation?	There is no evidence of any significant way in which any form of technology has been used to assist in weight estimation, whether through the use of mobile phone technology or the use of other high-tech methods to facilitate or augment weight estimation *Some work more recently on this subject.
Can IBW be estimated by existing weight estimation systems?	The manufacturers of the Broselow tape and the developers of the Traub-Johnson and Traub-Kichen formulas have claimed that these methods predict ideal body weight. This has never been formally evaluated, tested or published, however. There is no other published mention of methods to calculate IBW other than growth-chart based methods.
Usability of weight estimation systems	
How does weight estimation and drug dosing impact on cognitive stress and load during emergencies?	Although there is some speculative evidence about the cognitive stress associated with weight estimation, there have been no studies which have researched this directly.

Is there any benefit to having pre-calculated drug doses printed on the Broselow tape?	Although this has been assumed to be true by many experts, including the producers of the international advanced life support guidelines, there is no research-based evidence to support the value of the Broselow system at this stage.
How is the accuracy of weight estimation methods affected during real or simulated emergency care?	There have been no studies on this topic.
Is there a difference in the accuracy and time taken for weight estimation between cooperative and uncooperative patients?	There have been no studies on this topic.
Is there a difference in the accuracy and time taken for weight estimation between children in different clinically realistic positions?	There have been no studies on this topic.
Integration of weight estimation and drug dosing	
What drug dose deviation should be considered a dose error?	There is no internationally recognised evidence or consensus on this topic.
What is known about paediatric pharmacology and drug dosing in emergencies?	There is a reasonable amount of data on the perpetration of dosing calculation errors during emergencies, but very little information on how pharmacology might be different and how dosing should be modified.
Is there any need to calculate emergency drug doses differently in obese children?	Limited - There is good evidence outside the realm of emergency medicine on the correct dosing strategies for obese children, but almost no evidence exists to guide the administration of drugs during emergencies involving obese children
Do weight estimation and drug dosing apps work well in emergencies?	None - There have been no published studies in the use of mobile phone apps for emergency weight estimation and drug dosing.

Motivation for the thesis

This research was needed for a number of reasons, mostly because of the large gaps in the literature in this field, as described in the previous section. There were a few additional motivations for this research, arising from current widespread practices in paediatric emergency medicine that have not kept pace with recent evidence in the field. Although it has been well-documented that new innovations may take up to 17 years to be adopted in general medical practice, this delay may perpetuate potentially harmful aspects of medical care [240]. With particular reference to the provision of initial emergency care to children, and specifically procedures involving weight estimation and drug dosing, there are a number of suboptimal practices that could be identified through the publication of this research. This research could thus potentially contribute in a small way to an acceleration of the adoption and implementation of up-to-date evidence-based practice.

The specific practices that provided the motivation for this research were:

- Many healthcare providers still use inaccurate methods of weight estimation [241].
- Many healthcare providers still use weight estimation methods that are difficult to use and are detrimental to cognitive functioning during emergency medical care [63, 242].
- Many healthcare organisations still teach and advocate the use of inaccurate weight estimation methods [67].
- Many healthcare providers do not use appropriate resuscitation aids [43].
- Many healthcare providers do not use (or have access to) appropriate technology, especially in the form of mobile phone applications [232].

There were additional motivations to conduct the research, related to the need to refine and improve the performance of the PAWPER tape:

- How to increase the accuracy of weight estimation by the PAWPER tape in taller children and obese children.
- How to increase the accuracy and standardisation of habitus assessment for use with the PAWPER tape.

- To establish whether there is evidence of value in having limited data on pre-calculated drug doses printed on a tape. These findings would provide useful guidance about the design of future tape-based systems.

The specific problem statements relating to the research that was included in this thesis are described below.

2.17 Problem statements

2.17.1 Problem Statement 1

The performance of weight estimation systems in different populations has not been critically reviewed with the objective of establishing which methods are the most accurate and the most universally applicable (accurate in children of all ages, in a wide range of populations). This knowledge would help to establish which methods should be preferentially used in clinical practice.

2.17.2 Problem Statement 2

Many of the systems that do contain drug-dosing information, such as the Broselow tape, are portrayed as being self-sufficient but, in fact, are not. They all require supplementary reference material to provide sufficient information on drug dilution, preparation and administration to eliminate the need for additional – and error-ridden – calculations. The consequences of the inaccuracy of the Broselow tape as a weight estimation device and the incompleteness of the pre-calculated drug doses printed on the tape are potential delays and errors in drug dose calculations and drug delivery.

The evidence of the value of the Broselow tape system, in terms of the presentation of drug information, has never been systematically reviewed. This data is important to provide guidance on the development of future tape-based weight estimation systems: should they contain information on drug-dosing information, or should they be limited to weight-estimation information only? This data could also provide guidance on whether to encourage or discourage the use of length-based tapes with pre-calculated doses printed on the tape.

2.17.3 Problem Statement 3

The PAWPER tape performed extremely well in the initial alpha phase of testing and validation. However, two studies in obese populations in the USA showed that, while the PAWPER tape was still more accurate than any other weight estimation system tested, it was less accurate in obese children when compared with normal and underweight children. It was also not long enough to provide weight estimations in children taller than 153 cm. As a consequence, this could lead to drug dosing errors in obese children as well as the need to rely on another weight estimation method for taller children.

The newly created PAWPER XL tape would then need to undergo a preliminary validation study to evaluate its accuracy and its postulated superiority to the original PAWPER tape. Since many drugs need to be scaled to IBW rather than TBW in obese children, it would also be important to establish whether the new PAWPER XL system, designed to address its identified inadequacies, could estimate IBW in obese children.

2.17.4 Problem Statement 4

Simple gestalt visual inspection (anthroposcopy) is the easiest and quickest method of assessing a child's body habitus in order to assign a habitus score for PAWPER tape weight estimation. It has been verified to be accurate when used by both experienced and inexperienced users in South Africa. Although gestalt-based visual assessment of habitus has also previously been validated in other methodologies, it remains subjective, dependent on at least some experience and training as well as potentially difficult to generalise across populations as "average" in one population may not be the same as in another: what is "normal" in one population may be "overweight or obese" in another. Inexperienced users in international studies have experienced difficulties in estimating habitus in obese children.

Clarity on the best method of assessing children's habitus, within the time and cognitive constraints of an emergency care scenario would add to the strength of the PAWPER XL methodology: to establish whether the anthroposcopic method, a

checklist-based method or a reference image-based method would be most accurate and most accepted by users.

2.17.5 Problem Statement 5

Data from studies in the USA showed the original PAWPER system did not perform as well as expected in very obese populations and when used by inexperienced users. Failure to identify obesity in children and the resultant suboptimal assessment of habitus was the primary reason for the suboptimal weight estimation. This could lead to medication errors and the potential for adverse outcomes.

It is possible that these limitations could be improved by reducing the subjective nature of the habitus assessment by incorporating simple, proven anthropometric measurements into the PAWPER XL tape methodology. Measurements, such as mid-arm circumference, could be used to determine habitus in a more objective manner. This new model would need to be carefully evaluated in a preliminary database study to establish its accuracy, before proceeding to more extensive prospective studies.

2.17.6 Problem Statement 6

There is very limited data on the accuracy of weight estimation systems used during real or simulated emergencies. The ideal weight estimation system would perform as well when used during a real-life emergency as during non-emergency research conditions: for all sizes and ages of children, all medical and traumatic conditions, and in cooperative as well as uncooperative children in all physical positions. Whether the position of the child (e.g. supine vs. lateral recumbent or recovery position vs. seated) or degree of cooperation (cooperative vs. passively uncooperative vs. actively uncooperative) would influence the performance of each of these systems is also unknown, but probable. This could result in the failure to obtain a weight estimation, or a large error in weight estimation, both of which would severely negatively impact on the delivery of potentially lifesaving medical care.

It would therefore be useful to evaluate the performance of commonly-used weight estimation systems, such as the Broselow tape, the Mercy method and the PAWPER

XL tape under simulated emergency conditions. It would also be useful to evaluate a bespoke mobile phone weight-estimation application under the same conditions.

References

[1] Wells M, Goldstein L, Botha M. Emergency drug dosing in children. Edinburgh: Churchill Livingstone Elsevier; 2012.

[2] Kozer E, Scolnik D, Macpherson A, Keays T, Shi K, Luk T, et al. Variables associated with medication errors in paediatric emergency medicine. Pediatrics. 2002;110:737-742.

[3] Pan SD, Zhu LL, Chen M, Xia P, Zhou Q. Weight-based dosing in medication use: what should we know? Patient Preference and Adherence. 2016;10:549-60.

[4] Benjamin L, Frush K, Shaw K, Shook JE, Snow SK. Pediatric medication safety in the Emergency Department. Annals of Emergency Medicine. 2018;71(3):e17-e24.

[5] Rowe S, Siegel D, Benjamin DK, Jr., Best Pharmaceuticals for Children Act - Pediatric Trials Network Administrative Core C. Gaps in drug dosing for obese children: a systematic review of commonly prescribed emergency care medications. Clinical Therapeutics. 2015;37(9):1924-1932.

[6] Hansen M, Meckler G, Dickinson C, Dickenson K, Jui J, Lambert W, et al. Children's safety initiative: a national assessment of pediatric educational needs among emergency medical services providers. Prehosp Emerg Care. 2015;19(2):287-91.

[7] Hirata KM, Kang AH, Ramirez GV, Kimata C, Yamamoto LG. Pediatric weight errors and resultant medication dosing errors in the Emergency Department. Pediatric Emergency Care. 2017.

[8] Oakley P. Inaccuracy and delay in decision making in paediatric resuscitation, and a proposed reference chart to reduce error. British Medical Journal. 1988;297:817-819.

[9] Abdel-Rahman SM, Ridge A, Kearns GL. Estimation of body weight in children in the absence of scales: a necessary measurement to insure accurate drug dosing. Archives of Disease in Childhood. 2014;99(6):570-574.

[10] Lubitz DS, Seidel JS, Chameides L, Luten RC, Zaritsky AL, Campbell FW. A rapid method for estimating weight and resuscitation drug dosages from length in the pediatric age group. Annals of Emergency Medicine. 1988;17(6):576-581.

[11] Goldman J, Ojoo A, Abdel-Rahman S. Challenges in pediatric oral dosing. In: MacLeod S, Hill S, Koren G, Rane A, editors. Optimizing treatment for children in the developing world. Cham: Springer International Publishing; 2015. p. 33-44.

[12] Wells M, Goldstein L, Bentley A. The accuracy of emergency weight estimation systems in children - a systematic review and meta-analysis. International Journal of Emergency Medicine. 2017;10(29):1-43. https://doi.org/10.1186/s12245-017-0156-5.

[13] Young KD, Korotzer NC. Weight estimation methods in children: a systematic review. Annals of Emergency Medicine. 2016;68(4):441-451 e10.

[14] Garland JS, Kishaba RG, Nelson DB, Losek JD, Sobocinski KA. A rapid and accurate method of estimating body weight. American Journal of Emergency Medicine. 1986;4(5):390-393.

[15] Kozer E, Seto W, Verjee Z, Parshuram C, Khattak S, Koren G, et al. Prospective observational study on the incidence of medication errors during simulated resuscitation in a paediatric emergency department. British Medical Journal. 2004;329:1321-1325.

[16] The Joint Commission. Preventing paediatric medication errors 11 April 2008 [Accessed March 2018]. Issue 39:[Available from: http://www.jointcommission.org/SentinelEvents/SentinelEventAlert/sea_39.htm

[17] Paparella S. Weighing in on medication safety. Journal of Emergency Nursing. 2009;35(6):553-5.

[18] Kaufmann J, Roth B, Engelhardt T, Lechleuthner A, Laschat M, Hadamitzky C, et al. Development and prospective federal state-wide evaluation of a device for height-based dose recommendations in prehospital pediatric emergencies: a simple tool to prevent most severe drug errors. Prehospital Emergency Care. 2016;22(2):252-259.

[19] Claudet I, Bounes V, Federici S, Laporte E, Pajot C, Micheau P, et al. Epidemiology of admissions in a pediatric resuscitation room. Pediatr Emerg Care. 2009;25(5):312-6.

[20] Allie A. George EC paeds pts by triage. 2018.

[21] Yang W. Epidemiology of pediatric critically-ill patients presenting to the Pediatric Emergency Department. Klinische Pädiatrie. 2013;225:18-23.

[22] Christ M, Grossmann F, Winter D, Bingisser R, E P. Modern triage in the emergency department. Deutsches Ärzteblatt International. 2010;107(50):892-898.

[23] Ahmed M, Pai B, Reynolds T. Retention of knowledge of the Paediatric Life Support guidelines. Journal of the College of Physicians and Surgeons of Pakistan. 2012;22(3):194-5.

[24] O'Leary FM. Paediatric resuscitation training: is e-learning the answer? A before and after pilot study. J Paediatr Child Health. 2012;48(6):529-33.

[25] Luten R, Wears RL, Broselow J, Croskerry P, Joseph MM, Frush K. Managing the unique size-related issues of pediatric resuscitation: reducing cognitive load with resuscitation aids. Academic Emergency Medicine. 2002;9(8):840-7.

[26] Marlow RD, Wood DLB, Lyttle MD. Comparing the usability of paediatric weight estimation methods: a simulation study. Archives of Disease in Childhood. 2018.

[27] Weech A. Signposts on the highway of growth. American Journal of Diseases in Children. 1956;88:452-457.

[28] Kozer E, Scolnik D, Jarvis AD, Koren G. The effect of detection approaches on the reported incidence of tenfold errors. Drug Safety. 2006;29(2):169-74.

[29] Emergency care for children: growing pains. Washington, DC: National Academy Press, 2006.

[30] Barata IA, Benjamin LS, Mace SE, Herman MI, Goldman RD. Pediatric patient safety in the prehospital/emergency department setting. Pediatric Emergency Care. 2007;23(6):412-418.

[31] Mack A. Policy statement--principles of pediatric patient safety: reducing harm due to medical care. Pediatrics. 2011;127(6):1199-210.

[32] Croskerry P, Shapiro M, Campbell. S. Profiles in patient safety: medication errors in the emergency department. Academic Emergency Medicine. 2004;11:289-299.

[33] Selbst S, Levine S, Mull C. Preventing medical errors in pediatric emergency medicine. Pediatric Emergency Care. 2004;20:702-709.

[34] Frush KS. Medication errors in pediatric emergency care: developing a national standard? Annals of Emergency Medicine. 2007;50(4):369-370.

[35] Lesar TS, Mitchell A, Sommo P. Medication safety in critically ill children. Clinical Pediatric Emergency Medicine. 2006;7:215-225.

[36] Luten R. Error and time delay in pediatric trauma resuscitation: Addressing the problem with color-coded resuscitation aids. Surgical Clinics of North America. 2002;82(2):303-314.

[37] Conroy S, Appleby K, Rostock D, Unsworth V, Cousins D. Medication errors in a children's hospital. Paediatric and Perinatal Drug Therapy. 2007;8(1):18-25.

[38] Conroy S, Sweis D, Planner C, Yeung V, Collier J, Haines L, et al. Interventions to reduce dosing errors in children: A systematic review of the literature. Drug Safety. 2007;30(12):1111-1125.

[39] Rinke ML, Bundy DG, Velasquez CA, Rao S, Zerhouni Y, Lobner K, et al. Interventions to reduce pediatric medication errors: a systematic review. Pediatrics. 2014;134(2):338-60.

[40] Nichter MA. Medical errors affecting the paediatric intensive care patient: incidence, identification, and practical solutions. Pediatric Clinics of North America. 2008;55:757-777.

[41] Miller MR, Robinson KA, Lubomski LH, Rinke ML, Pronovost PJ. Medication errors in paediatric care: a systematic review of epidemiology and an evaluation of evidence supporting reduction strategy recommendations. Quality & Safety in Health Care. 2007;16(2):116-126.

[42] Ghaleb MA, Barber N, Franklin BD, Yeung VWS, Khaki ZF, Wong ICK. Systematic review of medication errors in pediatric patients. Annals of Pharmacotherapy. 2006;40(10):1766-1776.

[43] Kaufmann J, Laschat M, Wappler F. Medication errors in pediatric emergencies: a systematic analysis. Deutsches Ärzteblatt International. 2012;109(38):609-16.

[44] Lim CAE, Kaufman BJ, O'Connor Jr J, Cunningham SJ. Accuracy of weight estimates in pediatric patients by prehospital Emergency Medical Services personnel. American Journal of Emergency Medicine. 2013;31(7):1108-1112.

[45] Hoyle JD, Jr., Crowe RP, Bentley MA, Beltran G, Fales W. Pediatric prehospital medication dosing errors: a national survey of paramedics. Prehospital Emergency Care. 2017;21(2):185-191.

[46] Hoyle JD, Jr., Sleight D, Henry R, Chassee T, Fales B, Mavis B. Pediatric prehospital medication dosing errors: a mixed-methods study. Prehospital Emergency Care. 2015;20(1):117-124.

[47] Meckler G, Leonard J, Hoyle J. Pediatric patient safety in Emergency Medical Services. Clinical Pediatric Emergency Medicine. 2014;15(1):18-27.

[48] Marcin JP, Dharmar M, Cho M, Seifert LL, Cook JL, Cole SL, et al. Medication errors among acutely ill and injured children treated in rural emergency departments. Annals of Emergency Medicine. 2007;50:361-367.

[49] Srinivasan V, Nadkarni VM, Helfaer MA, Carey SM, Berg RA, American Heart Association National Registry of Cardiopulmonary Resuscitation I. Childhood obesity and survival after in-hospital pediatric cardiopulmonary resuscitation. Pediatrics. 2010;125(3):e481-8.

[50] Pinchevsky LE, Pesaturo KA, Smith BS, Hartman CA. Pilot comparison of three cardiopulmonary resuscitation medication dosing strategies in overweight children. Journal of Pediatric Pharmacology and Therapeutics. 2010;15(4):282-9.

[51] “accurate”. [Accessed 14 July 2018]. Available from: http://www.yourdictionary.com/accurate

[52] Stewart D. Accuracy of the Broselow tape for estimating paediatric weight in two Australian Emergency Departments: University of Sydney; 2009 [Accessed 27 September 2017, Master of Biostatistics]. Available from: https://ses.library.usyd.edu.au//bitstream/2123/6265/1/Declan%20Stewart%20WPP%20final.pdf

[53] Anderson M. Underestimation potentially better than overestimation. Archives of Disease in Childhood. 2007;92(10):937.

[54] Cattermole GN, Leung PY, Mak PS, Graham CA, Rainer TH. Mid-arm circumference can be used to estimate children's weights. Resuscitation. 2010;81(9):1105-10.

[55] Luten R, Zaritsky A. The sophistication of simplicity... optimizing emergency dosing. Academic Emergency Medicine. 2008;15(5):461-465.

[56] Cattermole GN, Leung MP, So HK, Mak PS, Graham CA, Rainer TH. Age-based formulae to estimate children's weight in the emergency department. Emergency Medicine Journal. 2011;28(5):390-396.

[57] Carley S. (@EMManchester). The other question is how much does it matter? If we are ‘out’ by 5/10/20/50 %?? At which point does it really matter? 29 May 2017 1:07pm. Tweet. Available from: https://twitter.com/EMManchester/status/869284162058027009

[58] McPhillips H, Stille C, Smith D, Pearson J, Stull J, Hecht J, et al. Methodological challenges in describing medication dosing errors in children. In: Henriksen K, Battles JB, Marks ES, Lewin DI, editors. Advances in Patient Safety: From Research to Implementation (Volume 2: Concepts and Methodology). Advances in Patient Safety. Rockville (MD)2005.

[59] Carasco CF, Fletcher P, Maconochie I. Review of commonly used age-based weight estimates for paediatric drug dosing in relation to the pharmacokinetic properties of resuscitation drugs. British Journal of Clinical Pharmacology. 2016;81(5):849-56.

[60] Kozer E, Scolnik D, MacPherson A, Rauchwerger D, Koren G. Using a preprinted order sheet to reduce prescription errors in a paediatric emergency department: a randomised, controlled trial. Pediatrics. 2005;116:1299-1302.

[61] Wells M, Goldstein L. Optimising emergency weight estimation in underweight and obese children: the accuracy of estimations of total body weight and ideal body weight. European Journal of Emergency Medicine. 2018; In Press.

[62] Murugan S, Parris P, Wells M. Drug preparation and administration errors during simulated paediatric resuscitations. Archives of Disease in Childhood. 2018; In Press.

[63] Winship C, Williams B, Boyle M. Which paediatric weight formula is best suited for the out-of-hospital field? Journal of Paramedic Practice. 2012;4(9):533-541.

[64] Ross EL, Heizer J, Mixon MA, Jorgensen J, Valdez CA, Czaja AS, et al. Development of recommendations for dosing of commonly prescribed medications in critically ill obese children. American Journal of Health-System Pharmacy. 2015;72(7):542-556.

[65] de Caen AR, Berg MD, Chameides L, Gooden CK, Hickey RW, Scott HF, et al. Part 12: Pediatric Advanced Life Support: 2015 American Heart Association guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care (Reprint). Pediatrics. 2015;136 Suppl 2:S176-95.

[66] Kleinman ME, Chameides L, Schexnayder SM, Samson RA, Hazinski MF, Atkins DL, et al. Pediatric advanced life support: 2010 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Pediatrics. 2010;126(5):e1361-99.

[67] Advanced Life Support Group. Advanced Paediatric Life Support: The Practical Approach (APLS) 6th Edition. New York: BMJ Books - Publisher: John Wiley & Sons (Wiley-Blackwell); 2016.

[68] American Heart Association. PALS provider manual. Dallas: American Heart Association; 2015.

[69] Nieman CT, Manacci CF, Super DM, Mancuso C, Fallon WF, Jr. Use of the Broselow tape may result in the underresuscitation of children. Academic Emergency Medicine. 2006;13(10):1011-9.

[70] Wells M, Coovadia A, Kramer E, Goldstein L. The PAWPER tape: A new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. Resuscitation. 2013;84(2):227-232.

[71] Abdel-Rahman SM, Paul IM, James LP, Lewandowski A. Evaluation of the Mercy TAPE: performance against the standard for pediatric weight estimation. Annals of Emergency Medicine. 2013;62(4):332-339.e6.

[72] Cattermole GN, Graham CA, Rainer TH. Age-based weight prediction is imprecise. Emergency Medicine Journal. 2011;29(1):82-83.

[73] Marlow R, Lo D, Walton L. Accurate paediatric weight estimation by age: mission impossible? Archives of Disease in Childhood. 2011;96(Suppl 1):A1-A2.

[74] Ralston ME, Myatt MA. Weight estimation tool for children aged 6 to 59 months in limited-resource settings. PLoS One. 2016;11(8):e0159260.

[75] Antevy P. Handtevy method helps providers rapidly calculate pediatric drug dosages 2013 [Accessed March 2018]. Available from: http://www.jems.com/articles/print/volume-38/issue-8/features/handtevy-method-helps-providers-rapidly.html

[76] So JL, Chow EP, Cattermole GN, Graham CA, Rainer TH. Finger counting method is more accurate than age-based weight estimation formulae in estimating the weight of Hong Kong children presenting to the emergency department. Emergency Medicine Australasia. 2016;28(6):691-697.

[77] Young TP, Chen BG, Kim TY, Thorp AW, Brown L. Finger counting: an alternative method for estimating pediatric weights. American Journal of Emergency Medicine. 2014;32(3):243-247.

[78] Young TP, Washington O, Flanery A, Guptill M, Reibling ET, Brown L, et al. Comparison of the finger counting method, the Broselow tape and common weight estimation formulae in Filipino children after Typhoon Haiyan. Emergency Medicine Australasia. 2015;27(3):239-244.

[79] Erker CG, Santamaria M, Moellmann M. Size does matter--age-related weight estimation in "tall n' thin" and "tiny n' thick" children and a new habitus-adapted alternative to the EPLS-formula. Resuscitation. 2014;85(9):1174-1178.

[80] Appelbaum N, Clarke J, Maconochie I, Darzi A. A model for habitus-adjusted paediatric weight estimation by age and data concerning the validation of this method on a large dataset of English children. Data Brief. 2018;16:771-774.

[81] Appelbaum N, Clarke J, Maconochie I, Darzi A. Paediatric weight estimation by age in the digital era: optimising a necessary evil. Resuscitation. 2018;122:29-35.

[82] Advanced Life Support Group. Advanced paediatric life support: the practical approach. 6e ed. Oxford: BMJ books / Blackwells; 2016.

[83] Lowe CG, Campwala RT, Ziv N, Wang VJ. The Broselow and Handtevy Resuscitation Tapes: a comparison of the performance of pediatric weight prediction. Prehospital and Disaster Medicine. 2016;31(4):364-75.

[84] Carley S. APLS weight estimation – don't do it (well almost never). 2013 [Accessed March 2018]. Available from: http://stemlynsblog.org/apls-estimation-formulas-do-not-safely-predict-weight-in-uk-children-st-emlyns/#comments

[85] Luscombe M. "Kid's aren't what they used to be": a study of paediatric patients' weights and their relationship to current weight estimation formulae. British Journal of Anaesthesia. 2005;95(4):578.

[86] Traub SL, Johnson CE. Comparison of methods of estimating creatinine clearance in children. American Journal of Hospital Pharmacy. 1980;37(2):195-201.

[87] Traub SL, Kichen L. Estimating ideal body mass in children. American Journal of Hospital Pharmacy. 1983;40(1):107-10.

[88] Black K, Barnett P, Wolfe R, Young S. Are methods used to estimate weight in children accurate? Emergency Medicine. 2002;14(2):160-165.

[89] Shirley P, Hooper M. Use of length-based and age-based calculations of weight in paediatric emergency care. Emergency Medicine Australasia. 2002;14:33.

[90] Varghese A, Vasudevan VK, Lewin S, Indumathi CK, Dinakar C, Subba Rao SD. Do the length-based (Broselow®) tape, APLS, Argall and Nelson's formulae accurately estimate weight of Indian children? Indian Pediatrics. 2006;43(10):889-894.

[91] Milne WK, Yasin A, Lubell R. Canadian children have outgrown the Broselow tape. Canadian Journal of Emergency Medicine. 2009;11:274.

[92] Clark MC, Lewis RJ, Fleischman RJ, Ogunniyi AA, Patel DS, Donaldson RI. Accuracy of the Broselow Tape in South Sudan, "The Hungriest Place on Earth". Academic Emergency Medicine. 2016;23(1):21-28.

[93] Losek JD, Garland J, Nelson DB. Body habitus + height = accurate weight estimate. Annals of Emergency Medicine. 1989;18(6):709.

[94] Britnell S, Taylor S, Koziol-McLain J. Emergency weight estimation lookup tables for New Zealand children aged 5-10 years. Emergency Medicine Australasia. 2016;28(5):558-63.

[95] Jung JY, Kwak YH, Kim DK, Suh D, Chang I, Yoon C, et al. An assessment of the accuracy of a novel weight estimation device for children. Emergecy Medicine Journal. 2017;34(3):163-169.

[96] Rappaport LD, Brou L, Givens T, Mandt M, Balakas A, Roswell K, et al. Comparison of errors using two length-based tape systems for prehospital care in children. Prehospital Emergency Care. 2016;20(4):508-17.

[97] Molyneux E, Brogan R, Mitchell G, Gove S. Children's weights: guess or measure by tape. Lancet. 1999;354:1616.

[98] Both CP, Schmitz A, Buehler PK, Weiss M, Schmidt AR. How accurate are pediatric emergency tapes? A comparison of 4 emergency tapes with different length-based weight categorization. Pediatric Emergency Care. 2017.

[99] Akabarian S, Vahedparast H, Ravanipour M, Mirzaei K. Comparison of Broselow tape measurements versus mother estimations of pediatric weights. Iranian South Medical Journal. 2013;16(2):138-143.

[100] Chaput JP, Katzmarzyk PT, Barnes JD, Fogelholm M, Hu G, Kuriyan R, et al. Mid-upper arm circumference as a screening tool for identifying children with obesity: a 12-country study. Pediatric Obesity. 2017;12(6):439-445.

[101] Fiorentino M, Sophonneary P, Laillou A, Whitney S, de Groot R, Perignon M, et al. Current MUAC cut-offs to screen for acute malnutrition need to be adapted to gender and age: the example of Cambodia. PLoS One. 2016;11(2):e0146442.

[102] Chomtho S, Fewtrell MS, Jaffe A, Williams JE, Wells JC. Evaluation of arm anthropometry for assessing pediatric body composition: evidence from healthy and sick children. Pediatric Research. 2006;59(6):860-5.

[103] Cattermole GN, Leung M, Mak P, So HK, Graham CA, Rainer TH. Children's weights correlate more strongly with mid-arm circumference (MAC) than with age, height or foot-length. Journal of Emergency Medicine. 2009;37(2):228-229.

[104] Wozniak R. The evaluation of potential weight-estimation methods in a primarily HIV positive cohort in Botswana for use in resource limited settings: The University of British Columbia; 2012.

[105] Abdel-Rahman SM, Ridge AL. An improved pediatric weight estimation strategy. Open Medical Devices Journal. 2012;4:87-97.

[106] Uesugi T, Okada N, Sakai K, Nishina K, Mikawa K, Shiga M. Accuracy of visual estimation of body height and weight in supine paediatric patients. Paediatric Anaesthesia. 2002;12(6):489-494.

[107] Partridge RL, Abramo TJ, Haggarty KA, Hearn R, Sutton KL, An AQ, et al. Analysis of parental and nurse weight estimates of children in the pediatric emergency department. Pediatric Emergency Care. 2009;25(12):816-8.

[108] Kelly A. Estimation of children's weight in medical emergencies. In: Preedy V, editor. Handbook of anthropometry: Physical measures of human form in health and disease. New York: Springer-Verlag; 2012. p. 1151-1161.

[109] Rosenberg M, Thundiyil J, Greenberger S, Rawal A, Latimer-Pierson J. Does physician estimates of pediatric patient weights lead to inaccurate medication dosages. Annals of Emergency Medicine. 2010;56(3 (Suppl September 2010)):S47.

[110] Goldman RD, Buskin S, Augarten A. Parental estimates of their child's weight: accurate for resuscitation drug doses. Pediatric Emergency Care. 1999;15(1):19-21.

[111] Harris M, Patterson J, Morse J. Doctors, nurses and parents are equally poor at estimating pediatric weights. Pediatric Emergency Care. 1999;15:17-18.

[112] O'Connor DP, Gugenheim JJ. Comparison of measured and parents' reported height and weight in children and adolescents. Obesity. 2011;19(5):1040-6.

[113] Nosaka N, Fujiwara T, Knaup E, Okada A, Tsukahara H. Validity of mothers' reports of children's weight in Japan. Acta Medica Okayama. 2016;70(4):255-9.

[114] Jackson J, Strauss CC, Lee AA, Hunter K. Parents' accuracy in estimating child weight status. Addictive Behaviors. 1990;15(1):65-8.

[115] Banach A, Wade TJ, Cairney J, Hay JA, Faught BE, O'Leary DD. Comparison of anthropometry and parent-reported height and weight among nine year olds. Canadian Journal of Public Health. 2007;98(4):251-253.

[116] Trakulsrichai S, Boonsri C, Chatchaipun P, Chunharas A. Accuracy of three methods used for Thai children's body weight estimation. Journal of the Medical Association of Thailand. 2012;95(9):1194-1199.

[117] Bavdekar SB, Sathe S, Jani P. Prediction of weight of Indian children aged up to two years based on foot-length: implications for emergency areas. Indian Pediatrics. 2006;43(2):125-130.

[118] Carroll W, Jay N, Alexander J. Towards better weight estimation in the seriously ill child - a comparison of methods. Archives of Disease in Childhood. 2001;84(Supp I):A12.

[119] Elgie LD, Williams AR. Using age on clothes size label to estimate weight in emergency paediatric patients. European Journal of Emergency Medicine. 2012;19(5):338-340.

[120] Haftel AJ, Khan N, Lev R, Schonfeld N. Hanging leg weight — a rapid technique for estimating total body weight in pediatric resuscitation. Annals of Emergency Medicine. 1990;19(5):523-526.

[121] Yamamoto LG, Inaba AS, Young LL. Improving length-based weight estimates by adding a body habitus (obesity) icon. Annals of Emergency Medicine. 2007;50(3, Supplement):S61-S62.

[122] Yamamoto LG, Inaba AS, Young LL, Anderson KM. Improving length-based weight estimates by adding a body habitus (obesity) icon. American Journal of Emergency Medicine. 2009;27(7):810-815.

[123] Abdel-Rahman SM, Ahlers N, Holmes A, Wright K, Harris A, Weigel J, et al. Validation of an improved pediatric weight estimation strategy. Journal of Pediatric Pharmacology and Therapeutics. 2013;18(2):112-121.

[124] Wells M. Weight prediction in children in the emergency department. Johannesburg: University of the Witwatersrand; 2009.

[125] Wells M. The PAWPER tape: a more accurate form of tape-based weight estimation. Sanguine. 2011;1(2):4 - 6.

[126] Wells M, Hodgson E. Ujuzi: practical pearl/perle pratique. African Journal of Emergency Medicine. 2012;2(1):33-34.

[127] Chavez H, Peterson R, Lo K, Arel M. Weight estimation in an inner city Pediatric Emergency Department: the effect of obesity. American Journal of Emergency Medicine. 2015;33(10):1364-1367.

[128] Garcia CM, Meltzer JA, Chan KN, Cunningham SJ. A validation study of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape-a new weight estimation tool. Journal of Pediatrics. 2015;167(1):173-177.e1.

[129] Goldstein L, Wells M. Fat is the new normal. Journal of Pediatrics. 2016;168:257.

[130] Wells M, Goldstein L, Bentley A. A validation study of the PAWPER XL tape: accurate estimation of both total and ideal body weight in children up to 16 years of age. Trauma and Emergency Care. 2017;2(4):1-8.

[131] Georgoulas V, Wells M. The PAWPER tape and the Mercy Method outperform other methods of weight estimation in children in South Africa. South African Medical Journal. 2016;106(9):933-939.

[132] O'Leary F, John B, Hann L, Pegiazoglou I, McGarvey K. Estimating weight in the Paediatric Emergency Department. Archives of Disease in Childhood. 2016; In press.

[133] Abdel-Rahman SM, Paul IM, Delmore P, James L, Fearn L, Atz A, et al. A weight estimation strategy for preterm and full-term infants. Global Pediatric Health. 2017;4:2333794X17748775.

[134] Abdel-Rahman SM, Jacobsen R, Watts JL, Doyle SL, O'Malley DM, Hefner TD, et al. Comparative performance of pediatric weight estimation techniques: a human factor errors analysis. Pediatric Emergency Care. 2017;33(8):548-552.

[135] Batmanabane G, Kumar Jena P, Dikshit R, Abdel-Rahman SM. Using the Mercy Method for weight estimation in Indian children. Global Pediatric Health. 2015; January-December:1-7.

[136] Dicko A, Alhousseini ML, Sidibé B, Traoré M, Abdel-Rahman SM. Evaluation of the Mercy weight estimation method in Ouelessebougou, Mali. BMC Public Health. 2014;14:270.

[137] Broselow tape 2011 edition A [package insert]. Armstrong Medical Inc. Lincolnshire, IL. USA. 2011.

[138] Ideal Clinic definitions, components and checklists. Department of Health, Republic of South Africa, 2018.

[139] Standard treatment guidelines and essential medicines list for South Africa: hospital level paediatrics. Department of Health, Republic of South Africa, 2017.

[140] Cheema B, Westwood A. Standards for Paediatric Emergency Care: Expert Consensus Report for Emergency Centres in Western Cape. 2015.

[141] Davies F, Gaushe-Hill M, Chu SY, Cheema B, Ang A, Caceres L, et al. International Standards of Care for Children in Emergency Departments. West Melbourne: INTERNATIONAL FEDERATION FOR EMERGENCY MEDICINE; 2012.

[142] K Shrestha, P Subedi, O Pandey, L Shakya, K Chhetri, House. D. Estimating the weight of children in Nepal by Broselow, PAWPER XL and Mercy method. World Journal of Emergency Medicine. 2018;9(4):276-281.

[143] Cattermole GN, Leung PY, Graham CA, Rainer TH. Too tall for the tape: the weight of schoolchildren who do not fit the Broselow tape. Emergency Medicine Journal. 2014;31(7):541-544.

[144] Milne K, Yasin A, Knight J, Noel D, Lubell R, Filler G. Ontario children have outgrown the Broselow tape. Canadian Journal of Emergency Medicine. 2012;14(1):25-30.

[145] Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 1999-2016.: National Health and Nutrition Examination Survey Data; 2016. Available from: https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx

[146] Cornelissen KK, Gledhill LJ, Cornelissen PL, Tovee MJ. Visual biases in judging body weight. British Journal of Health Psychology. 2016;21(3):555-69.

[147] Twarog JP, Politis MD, Woods EL, Daniel LM, Sonneville KR. Is obesity becoming the new normal? Age, gender and racial/ethnic differences in

parental misperception of obesity as being 'About the Right Weight'. International Journal of Obesity (London). 2016;40(7):1051-1055.

[148] Tanner JM. Reliability of anthroposcopic somatotyping. American Journal of Physical Anthropology. 1954;12(2):257-65.

[149] Munroe RA, Clarke HH, Heath BH. A somatotype method for young boys. American Journal of Physical Anthropology. 1969;30(2):195-202.

[150] Ciovirnache M, Popa M, Florea I, Ionescu V, Popescu H. Somatotyping in the obese children and adolescents. Endocrinologie. 1980;18(3):193-200.

[151] Monyeki KD, Toriola AL, de Ridder JH, Kemper HC, Steyn NP, Nthangeni ME, et al. Stability of somatotypes in 4 to 10 year-old rural South African girls. Annals of Human Biology. 2002;29(1):37-49.

[152] Somatotype and cardiovascular functions of rural South African children aged 6 to 13 years : Ellisras Longitudinal Growth and Health Study. Cardiovascular Journal of South Africa. 2005;16(2 Suppl):16.

[153] Genovese J. Can Body Mass Index (BMI) be used as a proxy for somatotype. The Social Science Journal. 2009;46:390-393.

[154] Maddan S, Walker J, Miller J. The BMI as a somatotypic measure of physique: a rejoinder to Jeremy EC Genovese. The Social Science Journal. 2009;46:394-401.

[155] Barker VC. The angle of obesity: a simple measurement for body typing. New Zealand Medical Journal. 1976;84(577):437-9.

[156] Lockner DW, Heyward VH, Griffin SE, Marques MB, Stolarczyk LM, Wagner DR. Cross-validation of modified fatness-specific bioelectrical impedance equations. International Journal of Sport Nutrition. 1999;9(1):48-59.

[157] Amador M, Rodriguez C. Testing of two score systems for the diagnosis of malnutrition. Acta Paediatrica, International Journal of Paediatrics. 1983;24(3):231-45.

[158] Summers L, Clingerman KJ, Yang X. Validation of a body condition scoring system in rhesus macaques (Macaca mulatta): assessment of body composition by using dual-energy X-ray absorptiometry. Journal of the American Association of Laboratory Animal Science. 2012;51(1):88-93.

[159] Pullan NB. Condition scoring of White Fulani cattle. Tropical Animal Health and Production. 1978;10(2):118-20.

[160] Russel A. Body condition scoring of sheep. In Practice. 1984;6(3):91-3.

[161] Carroll CL, Huntington PJ. Body condition scoring and weight estimation of horses. Equine Veterinary Journal. 1988;20(1):41-5.

[162] Ullman-Cullere MH, Foltz CJ. Body condition scoring: a rapid and accurate method for assessing health status in mice. Laboratory Animal Science. 1999;49(3):319-23.

[163] Easterly ME, Foltz CJ, Paulus MJ. Body condition scoring: comparing newly trained scorers and micro-computed tomography imaging. Laboratory Animals (NY). 2001;30(3):46-9.

[164] Dorsten CM, Cooper DM. Use of body condition scoring to manage body weight in dogs. Contemporary Topics in Laboratory Animal Science. 2004;43(3):34-7.

[165] Roche JR, Dillon PG, Stockdale CR, Baumgard LH, VanBaale MJ. Relationships among international body condition scoring systems. Journal of Dairy Science. 2004;87(9):3076-9.

[166] Clingerman KJ, Summers L. Development of a body condition scoring system for nonhuman primates using Macaca mulatta as a model. Laboratory Animals (NY). 2005;34(5):31-6.

[167] Bjornvad CR, Nielsen DH, Armstrong PJ, McEvoy F, Hoelmkjaer KM, Jensen KS, et al. Evaluation of a nine-point body condition scoring system in physically inactive pet cats. Am J Vet Res. 2011;72(4):433-7.

[168] Clingerman KJ, Summers L. Validation of a body condition scoring system in rhesus macaques (Macaca mulatta): inter- and intrarater variability. Journal of the American Association of Laboratory Animal Sciience. 2012;51(1):31-6.

[169] Dugdale AH, Grove-White D, Curtis GC, Harris PA, Argo CM. Body condition scoring as a predictor of body fat in horses and ponies. Veterinary Journal. 2012;194(2):173-8.

[170] Bercovich A, Edan Y, Alchanatis V, Moallem U, Parmet Y, Honig H, et al. Development of an automatic cow body condition scoring using body shape signature and Fourier descriptors. J Dairy Sci. 2013;96(12):8047-59.

[171] Vasseur E, Gibbons J, Rushen J, de Passille AM. Development and implementation of a training program to ensure high repeatability of body condition scoring of dairy cows. J Dairy Sci. 2013;96(7):4725-37.

[172] Isensee A, Leiber F, Bieber A, Spengler A, Ivemeyer S, Maurer V, et al. Comparison of a classical with a highly formularized body condition scoring system for dairy cattle. Animal. 2014;8(12):1971-7.

[173] Morfeld KA, Lehnhardt J, Alligood C, Bolling J, Brown JL. Development of a body condition scoring index for female African elephants validated by ultrasound measurements of subcutaneous fat. PLoS One. 2014;9(4):e93802.

[174] Vieira A, Brandao S, Monteiro A, Ajuda I, Stilwell G. Development and validation of a visual body condition scoring system for dairy goats with picture-based training. Journal of Dairy Science. 2015;98(9):6597-608.

[175] Wijeyamohan S, Treiber K, Schmitt D, Santiapillai C. A visual system for scoring body condition of Asian elephants (Elephas maximus). Zoo Biol. 2015;34(1):53-9.

[176] Spoliansky R, Edan Y, Parmet Y, Halachmi I. Development of automatic body condition scoring using a low-cost 3-dimensional Kinect camera. J Dairy Sci. 2016;99(9):7714-25.

[177] Bulik CM, Wade TD, Heath AC, Martin NG, Stunkard AJ, Eaves LJ. Relating body mass index to figural stimuli: population-based normative data for Caucasians. International Journal of Obesity and Related Metabolic Disorders. 2001;25(10):1517-24.

[178] Gardner RM, Brown D. Body image assessment: A review of figural drawing scales. Personality and Individual Differences. 2010;48:107-111.

[179] Collins M. Body figure perceptions and preferences among preadolescent children. International Journal of Eating Disorders. 1991;10(2):199-208.

[180] Truby H, Paxton SJ. Development of the Children's Body Image Scale. British Journal of Clinical Psychology. 2002;41(Pt 2):185-203.

[181] Yepes M, Viswanathan B, Bovet P, Maurer J. Validity of silhouette showcards as a measure of body size and obesity in a population in the African region: A practical research tool for general-purpose surveys. Population Health Metrics. 2015;13:35.

[182] Cortes-Martinez G, Vallejo-de la Cruz N, Perez-Salgado D, Ortiz-Hernandez L. Usefulness of body silhouettes to assess the nutritional status of school-age children and adolescents in Mexico City. Boletín Médico del Hospital Infantil de México. 2009;66:511-521.

[183] Hart RP, O'Shanick GJ. Forgetting rates for verbal, pictorial, and figural stimuli. Journal of Clinical and Experimental Neuropsychology. 1993;15(2):245-65.

[184] WHO Multicentre Growth Reference Study Group. Reliability of anthropometric measurements in the WHO Multicentre Growth Reference Study. Acta Paediatrica. 2006;450(Suppl):38-46.

[185] Hohenhaus SM, Frush KS. Pediatric patient safety: common problems in the use of resuscitative aids for simplifying pediatric emergency care. Journal of Emergency Nursing. 2004;30(1):49-51.

[186] Hohenhaus SM, Frush KS. Pediatric patient safety in resuscitation simulations: lessons learned from studying clinician behavior. Pediatric Emergency Care. 2005;21(1):71-75.

[187] Hunt EA, Hohenhaus SM, Luo X, Frush KS. Simulation of pediatric trauma stabilization in 35 North Carolina emergency departments: identification of targets for performance improvement. Pediatrics. 2006;117(3):641-648.

[188] Wong I. Ways to reduce drug dose calculation errors in children. Journal of Health Services Research & Policy. 2010;15 Suppl 1:68-70.

[189] Parshuram CS, To T, Seto W, Trope A, Koren G, Laupacis A. Systematic evaluation of errors occurring during the preparation of intravenous medication. Canadian Medical Association Journal. 2008;178(1):42-48.

[190] Pentin J, Smith J. Drug calculations: are they safer with or without a calculator? British Journal of Nursing. 2006;15(14):778-781.

[191] Hazinski MF. Reducing calculation errors in drug dosages: the pediatric critical information sheet. Pediatric Nursing. 1986;12(2):138-40.

[192] Potts M, Phelan K. Deficiencies on calculation and applied mathematics skills in pediatrics among primary care interns. Archives of Pediatric and Adolescent Medicine. 1996;150:748-752.

[193] Rowe C, Koren T, Koren G. Errors by paediatric residents in calculating drug doses. Archives of Disease in Childhood. 1998;79(1):56-8.

[194] Bernius M, Thibodeau B, Jones A, Clothier B, Witting M. Prevention of pediatric drug calculation errors by prehospital care providers. Prehospital Emergency Care. 2008;12(4):486-94.

[195] Avidan A, Levin PD, Weissman C, Gozal Y. Anesthesiologists' ability in calculating weight-based concentrations for pediatric drug infusions: an observational study. Journal of Clinical Anesthesia. 2014;26(4):276-80.

[196] Slishman S, Sapien R, Crandall CS. Introducing a simple, weight-based, color-coded, medication dosing device. Pediatric Emergency Care. 2002;18(3):212-215.

[197] Watkins SC, Anders S, Clebone A, Hughes E, Zeigler L, Patel V, et al. Paper or plastic? Simulation based evaluation of two versions of a cognitive aid for managing pediatric peri-operative critical events by anesthesia trainees: evaluation of the society for pediatric anesthesia emergency checklist. Journal of Clinical Monitoring and Computing. 2016;30(3):275-83.

[198] Marshall S. The use of cognitive aids during emergencies in anesthesia: a review of the literature. Anesthesia & Analgesia. 2013;117(5):1162-71.

[199] Burden AR, Carr ZJ, Staman GW, Littman JJ, Torjman MC. Does every code need a "reader?" improvement of rare event management with a cognitive aid "reader" during a simulated emergency: a pilot study. Simulation in Healthcare. 2012;7(1):1-9.

[200] Nelson KL, Shilkofski NA, Haggerty JA, Saliski M, Hunt EA. The use of cognitive aids during simulated pediatric cardiopulmonary arrests. Simulation in Healthcare. 2008;3(3):138-45.

[201] Lehmann CU, Kim GR. Using information technology to reduce pediatric medication errors. Journal of Clinical Outcomes Management. 2005;12(10):511-518.

[202] Kirk R, Li-Meng G, Packia J, Min Kam H, Ong B. Computer calculated dose in paediatric prescribing. Drug Safety. 2005;28(9):817-824.

[203] Menke JA, Broner CW, Campbell DY, McKissick MY, Edwards-Beckett JA. Computerized clinical documentation system in the pediatric intensive care unit. BMC Medical Informatics and Decision Making. 2001;1:3-10.

[204] Sexton J, Thomas E, Helmreich R. Error, stress, and teamwork in medicine and aviation: cross sectional surveys. British Medical Journal. 2000;320:745-749.

[205] Lancaster L. Mission impossible — minimising the terror of paediatric resuscitation for staff in the ED. Australasian Emergency Nursing Journal. 2002;5(3):16-20.

[206] Kahneman D. Thinking, fast and slow. New York: Farrar, Strauss and Giroux; 2011.

[207] Low D, Clark N, Soar J, Padkin A, Stoneham A, Perkins GD, et al. A randomised control trial to determine if use of the iResus(c) application on a smart phone improves the performance of an advanced life support provider in a simulated medical emergency. Anaesthesia. 2011;66(4):255-62.

[208] Bierbrier R, Lo V, Wu RC. Evaluation of the accuracy of smartphone medical calculation apps. Journal of Medical Internet Research. 2014;16(2):e32.

[209] Damhoff HN, Kuhn RJ, Baker-Justice SN. Medication preparation in pediatric emergencies: comparison of a web-based, standard-dose, bar code-enabled system and a traditional approach. Journal of Pediatric Pharmacological Therapy. 2014;19(3):174-81.

[210] Shannon T, Ratchford A, Southward D, Hildreth A. The development of a computerised equipment and drug calculator for use in resuscitation. Emergency Medicine Journal. 2002;19(3):215-218.

[211] Eppich WJ, Adler MD, McGaghie WC. Emergency and critical care pediatrics: use of medical simulation for training in acute pediatric emergencies. Current Opinion in Pediatrics. 2006;18(3):266-71.

[212] Eppich WJ, Brannen M, Hunt EA. Team training: implications for emergency and critical care pediatrics. Current Opinion in Pediatrics. 2008;20(3):255-60.

[213] Hohenhaus SM. Is this a drill? Improving pediatric emergency preparedness in North Carolina's emergency departments. Journal of Emergency Nursing. 2001;27(6):568-70.

[214] Lammers R, Byrwa M, Fales W. Root causes of errors in a simulated prehospital pediatric emergency. Academic Emergency Medicine. 2012;19(1):37-47.

[215] Porter E, Barcega B, Kim TY. Analysis of medication errors in simulated pediatric resuscitation by residents. Western Journal of Emergency Medicine. 2014;15(4):486-90.

[216] Gaca AM, Frush DP, Hohenhaus SM, Luo X, Ancarana A, Pickles A, et al. Enhancing pediatric safety: using simulation to assess radiology resident preparedness for anaphylaxis from intravenous contrast media. Radiology. 2007;245(1):236-44.

[217] Armstrong M, Lambert M, Sharwood K, Lambert E. Obesity and overweight in South African primary school children – the Health of the Nation Study. South African Medical Journal. 2006;96:439-444.

[218] Sturm R, Hattori A. Morbid obesity rates continue to rise rapidly in the United States. International Journal of Obesity (London). 2013;37(6):889-91.

[219] World Health Organisation. Double burden of malnutrition [Accessed 21 July 2018]. Available from: http://www.who.int/nutrition/double-burden-malnutrition/en/

[220] de Onis M, Blossner M, Borghi E. Global prevalence and trends of overweight and obesity among preschool children. American Journal of Clinical Nutrition. 2010;92(5):1257-64.

[221] Organisation WH. Overweight and obesity [Accessed 21 July 2018]. Available from: http://www.who.int/gho/ncd/risk_factors/overweight_obesity/overweight_adolescents/en/

[222] Brunette DD. Resuscitation of the morbidly obese patient. Emergency Medicine Journal. 2004;22:42-47.

[223] Brown CV, Neville AL, Salim A, Rhee P, Cologne K, Demetriades D. The impact of obesity on severely injured children and adolescents. Journal of Pediatric Surgery. 2006;41(1):88-91; discussion 88-91.

[224] Callaghan LC, Walker JD. An aid to drug dosing safety in obese children: development of a new nomogram and comparison with existing methods for estimation of ideal body weight and lean body mass. Anaesthesia. 2015;70(2):176-82.

[225] Harskamp-van Ginkel MW, Hill KD, Becker K, Testoni D, Cohen-Wolkowiez M, Gonzalez D, et al. Drug dosing and pharmacokinetics in children with obesity: a systematic review. Journal of the American Medical Association Pediatrics. 2015;169(7):678-85.

[226] Kendrick JG, Carr RR, Ensom MH. Pediatric obesity: pharmacokinetics and implications for drug dosing. Clinical Therapeutics. 2015;37(9):1897-923.

[227] Kendrick JG, Carr RR, Ensom MH. Pharmacokinetics and drug dosing in obese children. Journal of Pediatric Pharmacology and Therapeutics. 2010;15(2):94-109.

[228] Leykin Y, Miotto L, Pellis T. Pharmacokinetic considerations in the obese. Best Practice & Research Clinical Anaesthesiology. 2011;25(1):27-36.

[229] Dieckmann RA. The dilemma of paediatric drug dosing and equipment sizing in the era of patient safety. Emergency Medicine Australasia. 2007;19:490-493.

[230] Waters MR. Paediatric resuscitation. Dosage inaccuracies may be dangerous. British Medical Journal. 1993;306(6892):1612-3.

[231] Wells M, Kramer E. Optimizing emergency drug dosing in children. Academic Emergency Medicine. 2008;15(12):1325.

[232] Schmucker M, Heid J, Haag M. Development of an accommodative smartphone app for medical guidelines in pediatric emergencies. Studies in Health Technology and Informatics. 2014;198:87-92.

[233] Wetzel O, Schmidt AR, Seiler M, Scaramuzza D, Seifert B, Spahn DR, et al. A smartphone application to determine body length for body weight estimation in children: a prospective clinical trial. Journal of Clinical Monititoring and Computing. 2017.

[234] Xie B, Avila JI, Ng BK, Fan B, Loo V, Gilsanz V, et al. Accurate body composition measures from whole-body silhouettes. Medical Physics. 2015;42(8):4668-77.

[235] Morgan N, Luo X, Fortner C, Frush K. Opportunities for performance improvement in relation to medication administration during paediatric stabilisation. Quality & Safety in Health Care. 2006;15:179-183.

[236] Lerner C, Gaca AM, Frush DP, Hohenhaus S, Ancarana A, Seelinger TA, et al. Enhancing pediatric safety: assessing and improving resident competency in life-threatening events with a computer-based interactive resuscitation tool. Pediatric Radiology. 2009;39(7):703-9.

[237] Sandell JM, Charman SC. Can age-based estimates of weight be safely used when resuscitating children? Emergency Medicine Journal. 2009;26(1):43-7.

[238] Maconochie IK, Bingham R, Eich C, Lopez-Herce J, Rodriguez-Nunez A, Rajka T, et al. European Resuscitation Council Guidelines for Resuscitation 2015: Section 6. Paediatric life support. Resuscitation. 2015;95:223-48.

[239] de Caen AR, Berg MD, Chameides L, Gooden CK, Hickey RW, Scott HF, et al. Part 12: Pediatric Advanced Life Support: 2015 American Heart

Association guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation. 2015;132(18 Suppl 2):S526-42.

[240] Gawande A. The checklist manifesto: how to get things right. New York: Metropolitan Books; 2010.

[241] Wells M, Barnes L, Vincent-Lambert C. Paediatric weight estimation practices of advanced life support providers in Johannesburg, South Africa. African Journal of Emergency Medicine. 2018;8(2):51-54.

[242] Bowen L, Zyambo M, Snell D, Kinnear J, Bould MD. Evaluation of the accuracy of common weight estimation formulae in a Zambian paediatric surgical population. Anaesthesia. 2017;72(4):470-478.

Chapter 3

Aims and objectives

3.1 Aims in this thesis

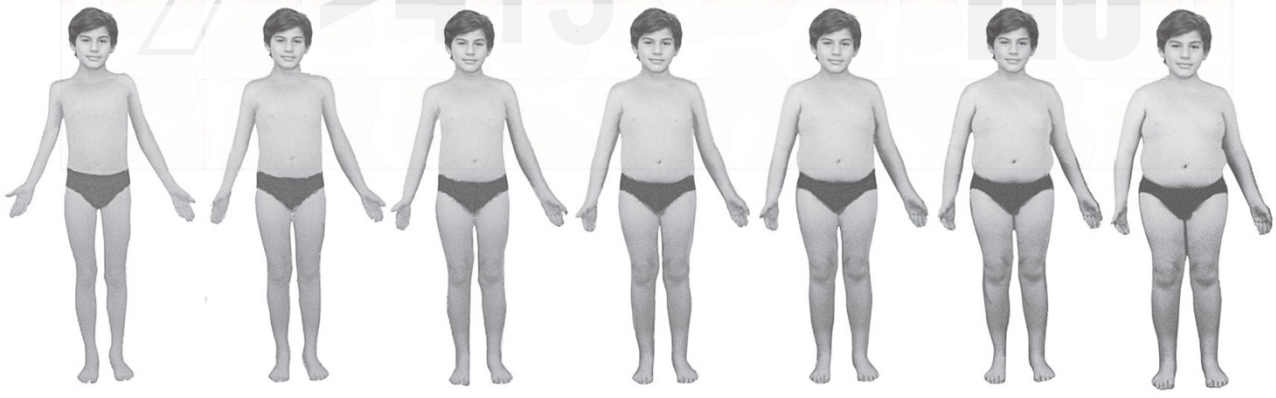
The overall aims of this project were to critically review the performance of existing weight estimation systems and, with the knowledge gained, to further advance and develop the PAWPER tape methodology.

3.2 Objectives

1. To describe the accuracy of commonly-used weight estimation systems in both high-income and low- and middle-income countries through the systematic review of previous studies and meta-analysis of data obtained from these studies.
2. To review the evidence supporting the inclusion of drug dosing information on length-based weight-estimation tapes, specifically the Broselow tape.
3. To develop a new version of or methodology for the PAWPER tape based on the findings of previous studies; to compare the weight estimation performance of this new version against other weight estimation systems; and to establish whether it could accurately predict both total and ideal body weight in obese children.
4. To establish whether the accuracy of weight estimation by the PAWPER XL method could be improved by the use of figural reference images to aid in the assessment of body habitus.
5. To develop and validate a new, standardised and more objective system of habitus assessment for the new PAWPER XL-MAC system using measurements of mid-arm circumference.
6. To evaluate the weight estimation accuracy of the PAWPER XL tape system, the EDD4C mobile phone application, the Mercy method and the Broselow tape during simulated paediatric emergencies.

Section 2

The accuracy of weight estimation systems



Chapter 4

The accuracy of emergency weight estimation systems in children

4.1 Prologue

4.1.1 Background

The first step when evaluating the functioning of any new methodology is to review the performance of existing similar systems to establish a benchmark standard. There are currently more than 40 different weight estimation systems that can be used in children and there have been more than 200 articles published on paediatric weight estimation. There is, however, as yet no established consensus on which weight estimation systems are the most accurate and no agreement on the standard of performance required of weight estimation systems in order to be adopted into clinical practice.

There has been only one older non-systematic review book chapter as well as one more recent systematic review published on weight estimation [1, 2]. In her book chapter on weight estimation during medical emergencies in children, published in 2012, Kelly suggested that parental estimates of weight, when they were available, were the most accurate method, followed by estimates using the Broselow tape [1]. In the more recent systematic review, published in 2016 by Young *et al*, the authors concluded that parent estimates and length- and habitus-based methods (such as the PAWPER tape and Mercy method) were the most accurate methods to predict children's total body weight [2]. The advent of the era of the modern dual length- and habitus-based methodologies, which began in 2012 with the first publications on the Mercy method and the PAWPER tape, had an immediate impact in the field of weight estimation [3, 4]. However, older, and less accurate methods, continue to be researched, used in clinical practice and taught in paediatric resuscitation courses [5].

No quantitative meta-analyses had yet been published on weight estimation. Meta-analysis allows for an aggregation of data which creates a greater statistical power and more robust assessment of accuracy than is possible from any individual research study. It thus provides a higher level of evidence than individual studies and a more solid foundation for the development of evidence-based practice.

This is the first systematic review with meta-analysis on the accuracy of weight estimation systems in children. A systematic review and meta-analysis focused on data from low- and middle-income countries was prepared and published at a similar time, but was not included in this thesis [6].

4.1.2 Purpose of the study

The primary purpose of this study was to determine the global accuracy of existing paediatric weight estimation systems. The problem was that there was just not enough conclusive information on the performance of each weight estimation system from individual studies because of variations between populations and differences in the statistical analyses presented. Despite many individual studies showing a poor accuracy for one-dimensional age- and length-based methods, for example, they continue to be researched with no fundamental change to the methodology to improve their accuracy. The meta-analysis would build evidence to allow users to make an informed, evidence-based decision about each system.

The other specific purposes of the meta-analysis included:

- Previous weight estimation reviews showed that parental estimates provided the most accurate weight estimates. It was essential to establish whether this held true when analysing pooled data, and whether a direct comparison between parental estimates and the newer dual length- and habitus-based systems could yield additional information.
- The newer, two-dimensional weight estimation systems have shown good accuracy in many different populations. It was therefore vital to evaluate the pooled data to confirm these findings to establish whether they were, in fact, globally and universally accurate. It was also important to compare these

systems with other methods to prove the superior performance found in individual studies.

- In terms of this PhD research, it was necessary to evaluate the PAWPER system, in comparison to other methodologies, to determine whether it was worth continuing with this methodology. The analysis would help to determine whether it needed to be kept unchanged, modified according to any evidence of short-comings or abandoned completely in favour of a more accurate and appropriate system.

The secondary purpose of this article was to investigate and establish a standard by which weight estimation systems could be judged. This would allow readers, researchers, reviewers and editors to decide whether a system had achieved a minimum level of accuracy to be worthy of publishing or using in clinical practice. It is a fundamental principle of statistical methodology to establish appropriate outcome criteria, which is almost universally absent in weight estimation studies.

This study provided the data to address Objective 1 (to describe the accuracy of commonly-used weight estimation systems in both high-income and low- and middle-income countries through the systematic review of previous studies and meta-analysis of data obtained from these studies).

4.2 Published paper

Wells M, Goldstein L, Bentley A. The accuracy of emergency weight estimation systems in children - a systematic review and meta-analysis. International Journal of Emergency Medicine 2017;10(29):1-43.

The PDF of the published article is included. The online-only supplementary material is appended after the PDF. The tables or figures have been retained in the formatting in which they are archived in the journal website.

Wells et al. International Journal of Emergency Medicine (2017) 10:29
DOI 10.1186/s12245-017-0156-5

International Journal of Emergency Medicine

REVIEW Open Access

The accuracy of emergency weight estimation systems in children—a systematic review and meta-analysis

Mike Wells[1,2*], Lara Nicole Goldstein[1] and Alison Bentley[1]

Abstract

The safe and effective administration of fluids and medications during the management of medical emergencies in children depends on an appropriately determined dose, based on body weight. Weight can often not be measured in these circumstances and a convenient, quick and accurate method of weight estimation is required. Most methods in current use are not accurate enough, but the newer length-based, habitus-modified (two-dimensional) systems have shown significantly higher accuracy. This meta-analysis evaluated the accuracy of weight estimation systems in children. Articles were screened for inclusion into two study arms: to determine an appropriate accuracy target for weight estimation systems; and to evaluate the accuracy of existing systems using standard meta-analysis techniques. There was no evidence found to support any specific goal of accuracy. Based on the findings of this study, a proposed minimum accuracy of 70% of estimations within 10% of actual weight (PW10 > 70%), and 95% within 20% of actual weight (PW20 > 95%) should be demonstrated by a weight estimation system before being considered to be accurate. In the meta-analysis, the two-dimensional systems performed best. The Mercy method (PW10 70.9%, PW20 95.3%), the PAWPER tape (PW10 78.0%, PW20 96.6%) and parental estimates (PW10 69.8%, PW20 87.1%) were the most accurate systems investigated, with the Broselow tape (PW10 55.6%, PW20 81.2%) achieving a lesser accuracy. Age-based estimates achieved a very low accuracy. Age- and length-based systems had a substantial difference in over- and underestimation of weight in high-income and low- and middle-income populations. A benchmark for minimum accuracy is recommended for weight estimation studies and a PW10 > 70% with PW20 > 95% is suggested. The Mercy method, the PAWPER tape and parental estimates were the most accurate weight estimation systems followed by length-based and age-based systems. The use of age-based formulas should be abandoned because of their poor accuracy.

Keywords: Weight estimation, Broselow tape, PAWPER tape, Mercy method

Introduction

It cannot be considered to be good medical practice to use a weight estimation system that is known to be inaccurate [1]. When children's weight cannot be measured during emergency care, an accurate, rapid estimation of weight is needed, as the safety and effectiveness of emergent interventions may ultimately depend on the accuracy of the weight estimation [2, 3]. Since most drug doses in children are based on weight, an accurate estimation of weight is important to ensure that a correct amount of medication is administered to achieve the desired effect, as well as to prevent the potential complications and side-effects of overdosing [4, 5]. This is relevant because most paediatric medication errors occur in the Emergency Department and most cases of resultant patient harm are related to incorrect dosing [6–8].

The problem is that most contemporary methods used to estimate children's weight have been shown to lack sufficient accuracy and consistency of performance in different populations [9]. Most existing weight estimation systems are "one-dimensional", because a single variable, usually age or length, is used in the weight estimation methodology. These systems fail because a single variable cannot adequately account for the biological variability of weight-for-age and weight-for-length

* Correspondence: mike.wells@emergencymedicine.co.za
[1]Division of Emergency Medicine, Faculty of Health Sciences, University of the Witwatersrand, 7 York Road, Parktown, Johannesburg 2193, South Africa
[2]Postnet Suite 429, Private Bag X1510, Glenvista 2058, South Africa

Springer Open

© The Author(s). 2017 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

[10, 11]. There is a wide variability of body habitus that is not accounted for in these weight-estimation systems, aggravated by the increasing levels of obesity affecting children [12, 13]. Newer, more promising, methods are the "two-dimensional" or dual length- and habitus-based systems, which include two variables in the estimation methodology: length (or a surrogate such as humerus or ulna length) and habitus (or a surrogate such as mid-arm circumference or waist circumference) [5, 14–17]. These have been shown to be much more accurate than the older, one-dimensional systems, in many studies [5, 15, 18–22].

Healthcare providers may also need more than one approach to emergency weight estimation: while parental estimates of weight can be very accurate, parents may not be present at the time that emergency care is required (especially in the prehospital environment) [9]. In these situations, an evidence-based alternative system may be required.

There has been a large amount of material published on weight estimation in children. It would be useful to combine the data from these studies to establish the accuracy of different methodologies both within and between different populations. Since many of the same weight estimation systems are used in populations with very different prevalences of underweight and obese children, it needs to be ascertained whether this impacts on the accuracy outcomes of these systems.

In order to create an evidence-based approach to emergency paediatric weight estimation, it is crucial to discover which methods predict weight most accurately and which are most appropriate for emergency use. This will enable clinicians to decide which systems they should incorporate into their clinical practice and will provide some guidance to those who administer, teach and train paediatric advanced life support on which systems are important.

The overall aim of this study was to determine which paediatric weight estimation systems most accurately estimate total body weight in children. The first objective was to determine whether there was evidence in the literature for an acceptable benchmark level of accuracy for a weight estimation system. The second objective was to extract and pool data on the performance of paediatric weight estimation systems to integrate the findings, provide a more comprehensive analysis on their functioning and identify those systems that operated best in diverse populations. The third objective was to directly compare the accuracy of paediatric weight estimation systems, for which paired data was available, using pooled data and meta-analysis techniques.

Only one meta-analysis has addressed this topic, but was limited to studies in low- and middle-income countries [23].

Methods

This systematic review and meta-analysis followed the PRISMA guidelines.

Search strategy

Online databases (MEDLINE, SCOPUS, Science Direct and Google) were interrogated for eligible studies, published between January 1983 and May 2017, using the following search terms: "paediatric weight estimation", "weight estimation children" and "Broselow tape". Citation lists of reviewed papers were examined for additional relevant articles. Studies in any language were included if English translations were obtainable. To minimise publication bias, all studies with adequate reporting were included, whether full-text articles, dissertations, abstracts, conference presentations or other unpublished data that had undergone some form of peer-review.

Study selection and eligibility criteria

All studies that evaluated weight-estimation methodologies were assessed for inclusion into the study by two separate investigators (MW and LG). Articles that contained discussions on desired targets of accuracy of weight estimation systems, or analysis of the performance of weight-estimation systems were included in the qualitative arm of the review. Studies that presented original data with either accuracy data (percentage of estimations within 10% of actual weight (PW10)) or bias and precision data (mean percentage error plus an appropriate indicator of variance), or both, were included in the meta-analysis. Studies that did not include original data, those that did not include usable data and those at high risk of bias (see below) were excluded from the meta-analysis (see Fig. 1).

Data abstraction and analysis

Data was extracted from the included studies independently by two researchers (MW, LG), cross-checked and confirmed. Standard statistics for meta-analysis of method-comparison studies were used [24], with an emphasis on evaluating accuracy (percentage of estimations within 10% of actual weight), bias (mean percentage error) as well as precision (limits of agreement of percentage error). Two methods of representing the pooled parametric and non-parametric data were employed: a fixed effects model weighted by inverse variance and a random effects model. In general, the random effects model was preferred because of the large variance within and between samples as well as the effects of several very large database studies that may have introduced bias.

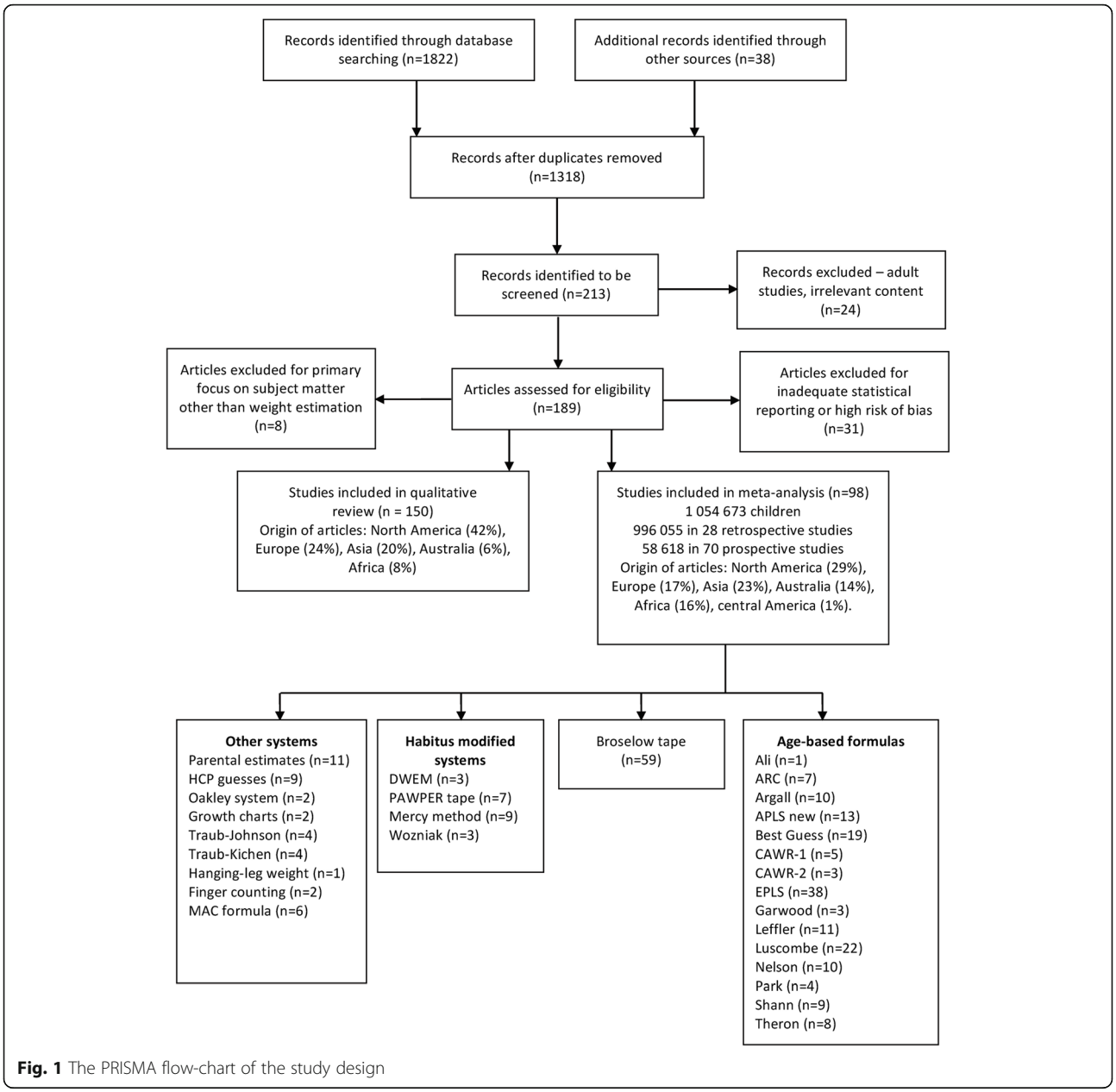


Fig. 1 The PRISMA flow-chart of the study design

Many of the evaluated studies presented incomplete data. Where it was possible, without risking bias, missing data was imputed using standard methodologies [25].

Direct comparisons between weight estimation systems, using pooled paired data, were performed with non-parametric techniques based on PW10 accuracy data, where such data was available.

Subgroup analysis

There was considerable heterogeneity in the use and composition of subgroups within the included studies. Wherever possible, subgroup analyses that had been performed in each study were included in the overall meta-analysis. The included subgroups focused on different age groups as previous studies have shown a difference in weight estimation accuracy between infants (<1 year), toddlers and pre-school children (1 to 6 years) and older children (>6 years of age) [26].

Risk of bias within and across studies

Reporting bias was minimised by including all available methodologically sound studies (published or not). Methodological causes of potential bias were common (e.g. the Broselow tape was not actually used in many studies, but weight-estimates were generated from length data), but these were individually assessed and

rated according to the level of risk of systematic bias. Studies with a high risk of bias were excluded from the meta-analysis (e.g. studies which excluded children above or below certain weight-for-length centiles).

Sensitivity analysis

There were three large database studies among those evaluated, with more than 100,000 children, one of which had more than 400,000 data points [27–29]. The effects of these "virtual" weight estimation studies, from very large databases, were carefully considered to establish any significant contribution to bias or distorted outcomes.

Software

Statistical analysis was performed using Stata (StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP), Graphpad Prism (GraphPad Prism version 8.00 for Mac, GraphPad Software, La Jolla, California, USA, www.graphpad.com) and Review manager (Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014).

Results

Excluded studies

The most common reason for exclusion of potentially relevant studies was incomplete data presentation (see Fig. 1). The large database studies did not have a significant impact on overall outcomes based on the sensitivity analysis and were therefore not excluded from the analysis.

Characteristics of included studies

Two-thirds of included studies evaluated multiple weight-estimation systems and contained paired data or made direct comparisons, while one-third evaluated only a single system. Prospective studies accounted for the majority of articles (70/98 (71.4%)) but a minority of total patients (58,618/1,054,673 (5.6%)).

Table 1 provides a descriptive summary of the studies included in both the qualitative review as well as the meta-analysis, including the major findings and limitations of each study and the risk of bias assessment for each included study.

Benchmark accuracy for a weight estimation system

After studying the 150 identified articles, only three articles were found to propose a statistically meaningful target for a weight estimation system: one article recommended that 95% of weight estimates must fall within 20% of actual weight and two articles suggested that 70% of estimates must be within 10% of actual weight *and* 95% of weight estimates must fall within 20% of actual weight [11, 30, 31]. There was, however, no evidence found upon which to base any specific measurement analysis metric for a weight estimation system. There was also no credible evidence found of a tolerable weight estimation error, in terms of safety for drug dose calculation, for an individual child.

In 90/150 articles (60.0%), there was no mention at all of an appropriate target for weight estimation accuracy. In 41/150 articles (27.3%) an error of < 10% was suggested as appropriate; in 11/150 articles (7.3%) an error of < 20% was advocated; in 2/150 articles (1.3%) an error of < 30%; and in 6/150 articles (4.0%) another value or a statistically inappropriate measure was proposed. None of the studies included any evidence to support these target figures. The values were selected based on clinical significance, pragmatic limits based on generalised therapeutic ratios, or based on guidelines on determining drug bioequivalence [32, 33].

Meta-analysis data on bias (trueness), precision and accuracy of paediatric weight estimation systems

Table 2 contains a description of each of the weight estimation systems reviewed, as well as any restrictions on their use. The raw data and outcomes for each of the weight-estimation methodologies included in the meta-analysis are shown in Additional file 1: Table S1. From the individual study data, it could be seen that there was very poor within-study precision for most weight estimation systems (shown by the wide limits of agreement), with the exception of the two-dimensional methods, which generally had precision limits of agreement of less than ± 20%.

Figure 2 shows the pooled data of the bias and precision for the weight-estimation systems evaluated. The fixed effects outcomes and data for the weight estimation methods not presented in Fig. 2 can be found in Table 3. The important findings can be summarised as follows:

- There was a wide variation in the weight estimation bias between low- and middle-income countries (overestimation) and high-income countries (underestimation). This was most noticeable with the age-based systems, less so with the length-based systems and least with the two-dimensional systems, which had virtually zero bias.
- There were very wide limits of agreement for all methods other than the PAWPER tape and the Mercy method.

Figure 3 show the overall accuracy data for each weight estimation system (PW10 data). Age-based systems were least accurate, length-based systems were slightly more accurate and parental estimates and the

Table 1 Studies included in the qualitative review and quantitative meta-analysis

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
Traub 1983 [69]	> 20,000	USA	R	0 to 18 years	Formula to estimate IBW (Traub-Kichen formula)	None	1	Low	Findings: Height was a good predictor of weight; IBW is only useful for a handful of drugs; TBW must be used in low weight-for-height children. Comments: Derivation study for Traub-Kichen formula. IBW predicted (actually 50th centile weight-for-length) by formula. Limitations: Incomplete presentation of data. Limited validation of formula.
Garland 1986 [67]	258	USA	P	0 to 19 years	DWEM, weight table	< 10%*	1, 2	Low	Findings: DWEM performed best of methods tested. Body habitus accurately assessed by evaluators. Comments: First ever report of evaluation of weight estimation systems in the literature. None of the systems tested were very accurate. Limitations: Only children up to 170 cm were included. Incomplete presentation of data.
Lubitz 1988 [70]	937	USA	P	0 to 12 years	Broselow tape	None	1, 2	Low	Findings: Broselow tape better than healthcare provider guesses and similar accuracy to DWEM. Accuracy of Broselow tape falls off sharply in children > 25 kg. Comments: Original study of Broselow tape. Authors recommended that an assessment of body habitus in children > 25 kg should be considered. Limitations: No formal, prospective comparison with other methodologies or indication of desired accuracy.
Oakley 1988 [71]	–	UK	–	–	–	None	1	N/A	Findings: Reference chart needed to aid rapid and accurate management. Comments: Weight estimation table derived from averaged boy-girl 50th centile weight-for-height (source not mentioned). Limitations: No validation of methodology.
Losek 1989 [72]	–	–	–	–	–	None	1	N/A	Findings: "Body habitus + height = accurate weight estimate". Comments: Letter claiming superior performance of the DWEM over the Broselow tape. No original data. Limitations: No mention of desired accuracy.
Haftel 1990 [73]	100	USA	P	2 months to 15 years	Hanging-leg weight	None	1, 2	Low	Findings: System accurate in children > 10 kg and more so > 25 kg. Comments: Good results never evaluated in subsequent studies. Limitations: Small sample size. Incomplete presentation of data.
Hughes 1990 [74]	139	UK	P	0 to 10 years	Broselow tape, healthcare provider guesses	None	1	N/A	Findings: Broselow tape performed substantially better than nurses' guesses. Comments: First validation study of Broselow tape in the UK. Limitations: Small sample size. Incomplete presentation of data.
Greig 1997 [36]	75	UK	P	0 to 12 years	healthcare provider guesses	None	1, 2	Low	Findings: Guesses of weight are very inaccurate; children should be weighed whenever possible. Comments: Authors suggest that accurate weight estimation is required for most drugs administered in emergency situations. Age-based formulas were wrongly considered acceptable. Limitations: Incomplete presentation of data, very small sample.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
Leffler 1997 [75]	117	USA	P	0 to 5 years	Parental estimates, Leffler	< 10%†	1, 2	Low	Findings: Parental estimates performed much better than formula. Distraught parents may be unreliable. Comments: Small sample size. Only children < 6 years included. Limitations: Over- or underestimation not recorded. Incomplete presentation of data.
Dearlove 1999 [76]	50	UK	P	1 to 16 years	Parental estimates, Broselow tape, EPLS, Argall	< 10%†	1, 2	Low	Findings: Broselow tape performed best, far better than parental estimates and age-based formulas. Comments: The target of 10% accuracy chosen for children was deliberately less than the 20% that the authors considered would be appropriate for adults. Limitations: Incomplete presentation of data and small sample size.
Goldman 1999 [77]	233	Israel	P	–	Parental estimates	< 10%†	1, 2	Low	Findings: Parents, especially mothers, can accurately estimate their children's weights. Comments: Those parents that had weighed their children an average of 5 weeks previously had the best results. The authors defined highly accurate weight estimations as < 5% error, accurate as < 10% error and semi-accurate as < 20% error. Limitations: Incomplete presentation of data. Misinterpret ation of bias as indicative of accuracy.
Harris 1999 [78]	100	USA	P	0 to 8 years	Parental estimates, healthcare provider guesses	None	1, 2	Low	Findings: Weight estimates by parents, nurses and doctors were significantly unreliable. The error is "so great and so frequent that clinically significant untoward effects can be anticipated". Comments: Broselow tape recommended by authors. Limitations: Incomplete presentation of data.
Molyneux 1999 [53]	142	Malawi	P	8 months to 5 years	Blantyre tape, healthcare provider guesses	<20%†	1, 2	Low	Findings: Healthcare provider guesses were very inaccurate; Blantyre tape better than guesses. A 20% error considered an acceptable target. Comments: Very young study population, mostly under 5 years. Limitations: Incomplete presentation of data.
Kun 2000 [79]	909	Hong Kong	P	0 to 12 years	Broselow tape	<10%*	1, 2	Low	Findings: Broselow tape most accurate in children from 10 to 25 kg, but acceptable for all children. Adjustment for habitus would be advantageous. Comments: Accuracy of Broselow tape outside of the 10–25 kg range was actually poor. The accuracy in this range was reasonable, but not as good as the authors suggest. Limitations: Poor interpretation of statistics. Broselow tape version not reported.
Carroll 2001 [80]	169	UK	P	–	EPLS, novel methods	None	1	N/A	Findings: MAC and shoe size were better indicators of weight than age. Comments: Abstract. Interesting concept, frequently cited abstract. Limitations: Incomplete presentation of data.
Vilke 2001 [81]	80	USA	P	–	Broselow tape, healthcare provider guesses	<50%	1	N/A	Findings: 95% of estimates within acceptable error range. Comments: Unrealistic target range, with no evidence basis. Tenfold errors in drug doses in 10% of cases;

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
									Broselow tape more accurate than guesses. Limitations: Incomplete presentation of data.
Black 2002 [52]	495	Australia	P	0 to 18 years	EPLS, Broselow tape, DWEM, Oakley, TJ, TK	None	1, 2	Low	Findings: Broselow tape and DWEM were more accurate than formulas. These methods should be used if weighing not possible. Comments: EPLS worst performer but poor accuracy of all systems. Good reproducibility of assessment of body habitus. Limitations: Incomplete presentation of data. Broselow tape version not reported.
Hofer 2002 [82]	585	Switzerland	R	6 months to 11 years	Broselow tape	<10%*	1, 2	Low	Findings: Broselow tape was accurate but underestimated weight in older children. Comments: Nearly 25% of sample excluded because they were too tall for the tape. Limitations: Broselow tape not actually used and version not reported. Incomplete presentation of data.
Uesegi 2002 [83]	48	Japan	P	–	Healthcare provider guesses	<20%	1, 2	Low	Findings: Doctors' guesses of children's weight were not accurate—drug doses should therefore be titrated in small paediatric patients. Comments: Wide variation in different doctors' accuracy, not related to seniority. All estimators were very inaccurate; worst estimations occurred in children < 20 kg. Limitations: Incomplete presentation of data; conclusion that underestimation of weight may "not be a serious problem" was not supported by the evidence.
Argall 2003 [84]	300	UK	P	1 to 10 years	EPLS, Broselow tape	None	1	N/A	Findings: Both methods performed poorly and worsened with increasing age. Comments: Difficult to draw any conclusions from this study, but Broselow tape marginally better than formula. Authors suggest that methods of weight estimation not keeping up with increasing obesity. Limitations: Broselow tape version not reported. Incomplete presentation of data.
Potier 2003 [85]	–	–	–	–	EPLS, Broselow tape	None	1	N/A	Findings: EPLS "may" be losing accuracy with increasing obesity; Broselow tape "may indeed be" more accurate. Comments: Mini-PICO analysis. Limitations: Limited qualitative-only evaluation. No comment on acceptable degree of accuracy.
Hohenhaus 2004 [86]	–	–	–	–	Broselow tape	None	1	N/A	Findings: Broselow tape may cause significant weight estimation errors if used incorrectly. Broselow tape more intended for equipment size determination than for weight estimation. Limitations: Broselow tape proposed as best instrument with minimal discussion. No targets for weight estimation.
Moore 2004 [87]	144	USA	P	–	Broselow tape	None	1	N/A	Findings: Despite using Broselow tape, only 56% correct medication doses delivered. Comments: Despite improving weight estimation, the Broselow tape did not decrease medication errors. Limitations: Incomplete presentation of data.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
Luscombe 2005 [88]	237	UK and New Zealand	P	1 to 10 years	EPLS, Leffler	None	1, 2	Low	Findings: The Leffler formula was more accurate than the EPLS formula. Comments: A new formula was promoted but the evidence was limited: only bias was evaluated, not accuracy. Both formulas underestimated weight significantly The Luscombe formula, as known today, was the result of a later study. Limitations: Incomplete presentation of data.
Theron 2005 [89]	909	New Zealand	P	1 to 10 years	EPLS, Shann, Leffler, Oakley, Broselow tape	None	1, 2	Low	Findings: Formulas and Broselow tape underestimated the weight of Pacific Island and Maori children. Comments: Broselow tape was the best performer of the systems tested, despite the fact that the authors reported the contrary. Limitations: Broselow tape version not reported. Incomplete presentation of data.
Bavdekar 2006 [90]	500	India	P	0 to 2 years	Novel formula based on foot length	None	1	N/A	Findings: Foot length can be used for emergency drug calculation Comments: Only 10% of population > 1 year old. very poor statistics. Method of data analysis made findings unreliable and uninterpretable - bias confused with accuracy.
Nieman 2006 [37]	7813	USA	R	0 to 12 years	Broselow tape	< 10%*	1, 2	Low	Findings: One third of children had inaccurate weight estimations. The 1998 tape performed better than the 2002 tape. A measurement of obesity (e.g. MAC) should be added to Broselow tape to increase accuracy. Comments: Broselow tape version 1998 and 2002A used. The study population BMI was 16.8. Good statistics. Limitations: Broselow tape not actually used. Nearly 7% of sample excluded because too tall for the tape.
Varghese 2006 [91]	500	India	P	1 to 12 years	Argall, EPLS, Nelson, Broselow tape	None	1	N/A	Findings: Formulas overestimated weight in this developing-world study. The Broselow tape was the most accurate. Comments: More than half the study population was under 6 months of age and only 8% > 5 years. Limitations: Incom plete presentation of data.
Anderson 2007 [92]	–	–	–	–	–	None	1	N/A	Comments: Letter. The author comments on benefits of under- vs. overestimation of weight. No quantification of permissible or safe error was suggested.
Dieckman 2007 [93]	–	–	–	–	–	None	1	N/A	Comments: Editorial on weight estimation and drug dosing. LBW might be better than TBW. No evidence offered for opinion; no weight estimation target suggested (for TBW or LBW).
Du Bois 2007 [94]	400	USA	P	0 to 19 years	Broselow tape, DWEM	MPE < 5%	1, 2	Low	Findings: DWEM better than Broselow tape, but both systems underestimated weight especially in children > 20 kg. Comments: Good inter-rater reliability for habitus assessment. Performance of DWEM much better than Broselow tape, but both systems performed poorly. Limitations: Incomplete presentation of data. Broselow tape version not reported. Statistically inappropriate accuracy target.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
Hashikawa 2007 [32]	1207	USA	R	0 to 12 years	Broselow tape	< 20%†	1	N/A	Findings: Approximately 60% accuracy of colour zones assignment (accurate drug dosing). Weight underestimated in obese and older children. Comments: Rising prevalence of obesity blamed for poor performance. Broselow tape version 2002B. The average study population BMI was 17. Limitations: Broselow tape not actually used. Assessment of correct zone assignment only, weight not measured. Incomplete presentation of data.
Im 2007 [95]	454	Korea	P	–	Broselow tape	None	1	High	Findings: Broselow tape accurate in children of normal weight-for-length Comments: Only children falling within 3rd to 97th weight-for-height centiles included. Very young study population. Broselow tape only recommended by authors for "normal-growth" children < 20 kg and < 120 cm. Limitations: Incomplete presentation of data. Conclusions not supported by findings.
Jang 2007 [96]	665	Korea	R	–	Broselow tape	< 10%	1, 2	Low	Findings: Broselow tape reasonably accurate in this population, but less so in children > 25 kg. Comments: Overall underestimation of weight. Performance not very good and on par with most other studies. Limitations: Broselow tape not actually used and version not reported. Incomplete presentation of data.
Kelly 2007 [97]	410	Australia	P	1 to 11 years	BG	< 20%	1, 2	Low	Findings: BG performed moderately well, but overestimated weight in low BMI children. Comments: Multiple papers on same data. BMI was 17 in study population. Significant number of children had large errors of weight estimation. Limitations: Incomplete presentation of data.
Krieser 2007 [98]	410	Australia	P	1 to 10 years	Parental estimates, Broselow tape, BG, Argall, EPLS	< 10%*	1, 2	Low	Findings: Parental estimates performed best, followed by Broselow tape. Only 11% of parents could not provide an estimate. Formulas performed much worse than other methods. Comments: Multiple papers on same data. Study population BMI was 17.1. Limitations: Broselow tape not actually used. Broselow tape version not reported. Incomplete presentation of data.
Luscombe 2007 [1]	13,988	UK	R	1 to 10 years	EPLS, Luscombe	None	1, 2	Low	Findings: The authors commented that since few children with high-acuity conditions are actually weighed in clinical practice, weight estimation essential. The EPLS formula significantly underestimated weight, which may lead to under-resuscitation. The Luscombe formula was more accurate. Comments: Both formulas actually performed poorly. Limitations: Incomplete presentation of data. Mean bias used incorrectly.
Luten 2007 [99]	–	–	–	–	Broselow tape	None	1	N/A	Findings: No substantiation for setting acceptable accuracy of weight estimation at 10%. Comments: Editorial comment; no evidence provided.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
Nguyen 2007 [100]	410	Australia	P	1 to 11 years	Argall	< 10%*	1, 2	Low	Findings: Argall formula performed poorly, especially in children > 35 kg. Comments: Study population BMI was 17. Limitations: Incomplete presentation of data.
Patel 2007 [101]	360	USA	P	1 month to 10 years	Parental estimates, EPLS, Broselow tape	None	1	N/A	Findings: Parental estimates performed best, then Broselow tape, then EPLS. Comments: Abstract. Conclusions based on correlation and rudimentary statistical analysis. Limitations: Incomplete presentation of data.
Pollock 2007 [48]	100	Malawi	P	1 to 7 years	EPLS, Luscombe	None	1, 2	Low	Findings: The Luscombe formula was less accurate than EPLS with greater overestimation of weight. The authors suggested length-based systems should rather be used. Comments: Scientific letter. Both formulas performed very poorly with significant overestimation of weight. Limitations: Incomplete presentation of data.
Thompson 2007 [34]	1843	Australia	R	0 to 14 years	BG, EPLS, ARC	None	1, 2	Low	Findings: The BG formula performed better than EPLS and ARC formulas. Authors advised cautious use in infants. Comments: Data obtained from high acuity patients. The BG did not actually perform that well in this study. Limitations: Incomplete presentation of data. Only measures of bias used for comparison.
Tinning 2007 [102]	67,363	Australia	R	0 to 14 years	BG development study	None	1	N/A	Findings: New formulas developed with no target accuracy and no validation sample. Comments: Limited indication of accuracy of formulas as only measures of bias reported. Limitations: Incomplete presentation of data.
Zikos 2007 [103]									Comments: Identical abstract and data to Patel 2007
Gardner 2008 [104]	not reported	UK	P	1 to 15 years	EPLS, Luscombe, healthcare provider guesses	None	1	N/A	Findings: Age-based formulas often used as basis for weight guess, but the formulas were better than guesses. Comments: Abstract. Limitations: Incomplete presentation of data.
Luten 2008 [3]	–	–	–	–	–	None	1	N/A	Comments: This was a commentary on Broselow tape—that IBW might be better than TBW as a goal; "the Broselow tape is a tool that was not designed to be used without clinical judgement". No evidence offered to support opinion; no goal target suggested (for TBW or IBW).
Ramarajan 2008 [56]	548	India	P	0 to 12 years	Broselow tape	< 10%†	1, 2	Low	Findings: Broselow tape overestimated weight by > 10% in Indian children over 10 kg. A correction factor was developed, but not validated. Comments: One of the few studies to show overestimation of weight by Broselow tape. Limitations: Broselow tape version not reported. Incomplete presentation of data.
Wells 2008 [105]	–	–	–	–	–	None	1	N/A	Comments: Letter. The Broselow tape has not been shown to accurately predict IBW. TBW should be used for dose calculation in emergencies.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
Zink 2008 [106]	127	USA	P	0 to 17 years	Parental estimates, healthcare provider guesses, DWEM, Broselow tape	None	1	High	Findings: Broselow tape and DWEM were the least accurate methods. Comments: No conclusions can be drawn from this study because of the methodology. Limitations: The data appears to favour healthcare provider and parent guesses, but the statistical methodology is flawed.
Anstett 2009 [107]	545	Ireland	R	–	Broselow tape	None	1, 2	Low	Findings: The Broselow tape was often inaccurate and tended to underestimate weight. Comments: Abstract. The Broselow tape actually performed better in this study than in many other studies. Limitations: Broselow tape version not reported. Broselow tape not actually used.
Cattamanchi 2009 [108]	15,000	India	P	2 months to 12 years	Broselow tape	< 10%*	1, 2	Low	Findings: The Broselow tape performed well, especially in children < 10 kg but underestimated all others, especially in children > 18 kg. Comments: Abstract. Very large prospective study. The authors recommended a new version of Broselow tape for Indian children because of underestimation of weight. Limitations: Broselow tape version not reported. Incomplete presentation of data.
Cattermole 2009 [109]	1368	Hong Kong	P	1 to 12 years	MAC, Broselow tape, foot length	None	1	N/A	Findings: Estimates of weight can be based on MAC. A special colour-coded MAC tape could be produced to aid drug dosing. The authors recommended habitus modified use of Broselow tape. Comments: Abstract. No data presented. Broselow tape performed better in younger children, MAC better in older children. Limitations: No data presentation.
Partridge 2009 [110]	777	USA	P	0 to 20 years	Parental estimates, healthcare provider guesses	< 10%†	1, 2	Low	Findings: Parents were better than nurses at estimating weight; nurses were very inaccurate. Comments: Guessed weights most often underestimations. The longer the time from last weighing, the greater the error. All nurses, regardless of training and experience, were poor estimators. Limitations: Incomplete presentation of data.
Paw 2009 [111]	791	UK	P	1 to 12 years	EPLS	None	1	High	Findings: Very poor performance of EPLS formula. The authors recommended the Broselow tape or an alternative formula. Comments: Abstract. Uniformly abysmal accuracy across different ethnic groups. Limitations: Incomplete presentation of data.
Sandell 2009 [112]	846	UK	P	1 to 11 years	EPLS, age-based estimates vs. length-based estimates	None	1, 2	Low	Findings: Length-based and age-based systems are suit able in emergencies, but length-based were better; new formulas more accurate than EPLS; "one size fits all" approach not likely to be successful. Comments: Unique method of analysing data—does not allow comparisons with other studies in this format. Age-based methods less accurate than suggested; biological variability less in length-based than age-based systems. Limitations: Incomplete presentation of data.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
So 2009 [113]	1011	USA	P	0 to 10 years	Broselow tape, Leffler, Theron	None	1, 2	Low	Findings: Broselow tape most accurate weight estimation method with best accuracy in normal BMI children. Comments: Study population BMI was 17.8. New formula developed, but not tested. Limitations: Incomplete presentation of data. Bias mistaken for accuracy. Broselow tape not actually used and Broselow tape version not reported.
Stewart 2009 [30]	475	Australia	P	0 to 10 years	Broselow tape	< 20%[†]	1, 2	Low	Findings: Best performance of Broselow tape between 10 and 25 kg but very inaccurate in children >25 kg. More accurate than age-based formulas, however. Comments: Masters dissertation. Good statistics. Limitations: Broselow tape version not reported.
Yamamoto 2009 [12]	542	Hawaii	P	–	Broselow tape, novel habitus-specific equation	None	1	High	Findings: Adding body habitus data to length increases accuracy of weight estimation. Comments: Useful data unable to be imputed from poorly presented statistics; overweight children were preferentially selected in this study Limitations: Incomplete presentation of data. Broselow tape version not reported and Broselow tape not actually used.
Bicer 2010 [114]	2319	Turkey	P	3 to 18 years	EPLS, Argall, novel formulas	None	1	High	Findings: Argall formula performed best. Comments: Article in Turkish; many formulas studied, often applied outside of the age-range for which they were intended. Limitations: Incomplete presentation of data. Measures of bias mistaken for measures of accuracy.
Casey 2010 [115]	1426	Australia	P	0 to 14 years	EPLS, BG, Broselow tape	< 20%[†]	1, 2	Low	Findings: BG was accurate in children 1–4 years of age. The EPLS formula was the least accurate of all methods. The authors recommend that the accuracy and ease-of-use of the Broselow tape mandates its use in the ED and pre-hospital. Comments: The authors' suggestions that the BG was more "accurate" than the Broselow tape are misleading: the Broselow tape actually significantly outperformed the BG formula. BMI was 17.9 in study population. Limitations: I ncomplete presentation of data. Broselow tape not actually used and Broselow tape version not reported.
Cattermole 2010 [116]	1370	Hong Kong	P	0 to 11 years	EPLS, Broselow tape, MAC formula	20–30%	1, 2	Low	Findings: MAC formula outperformed Broselow tape in children > 5 years. Comments: MAC actually only more accurate in children from 9 to 11 years who were too tall for the Broselow tape. Poor accuracy of all systems demonstrated. Limitations: Broselow tape not actually used, and 1998 edition values used. No evidence provided for desired accuracy targets.
Lulic 2010 [117]	209	Croatia	P	0 to 14 years	EPLS, Argall, BG, Luscombe, Broselow tape	< 10%[†]	1, 2	Low	Findings: Broselow tape was more accurate than age-based formulas but tended to slightly underestimate weight. Comments: Abstract; none of the systems estimated weight with an acceptable degree of accuracy. Limitations: Broselow tape version not reported. Incomplete presentation of data.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
Rosenberg 2010 [118]	401	USA	P	0 to 14 years	healthcare provider guesses	None	1	Low	Findings: Large errors in adrenaline doses were possible, based on doctors' guesses of weight. Comments: Abstract. Poor weight estimates translated into potentially harmful medication errors. Limitations: Incomplete presentation of data.
Williams 2010 [35]	468	Australia	P	4 and 6 years	healthcare provider guesses	None	1, 2	Low	Findings: Both age and weight were estimated poorly by paramedics, with no association with experience or training. Comments: Estimation of weight and age from images of two children by 234 estimators. Limitations: Incomplete presentation of data.
Bourdeau 2011 [119]	243	Canada	R	1 to 10 years	Broselow tape	< 10%†	1, 2	Low	Findings: Broselow tape was not accurate in First Nations' children and may need to increase weight estimates by 12%. Although ideal method of dosing is still unknown, it is best practice to eliminate any possible contributing inaccuracies. Comments: First Nations' children selected with high incidence of obesity in this population. Broselow tape did not perform equally well in all ethnic/population groups. Limitations: Incorrect use of some statistics. Broselow tape not actually used and Broselow tape version not reported.
Cattermole 2011 [45]	1248	Hong Kong	P	1 to 10 years	EPLS, ARC, Nelson, Shann, Leffler, BG, Argall, Luscombe, Theron, CAWR-1, CAWR-2	< 30%	1, 2	Low	Findings: CAWR-1 should be used in younger Chinese children, but other methods of weight estimation should be used in older children as age formulas are inaccurate. Comments: Comprehensive statistics. Findings suggested that no age formula performed well at any age. Limitations: No evidence to support suggested target of weight estimation.
Cattermole 2011 [120]	–	–	–	–	–	< 40%	1	N/A	Comments: Letter. "Age-based paediatric weight estimation is imprecise". The author reflects that an error of ± 40% might not be considered acceptable by most clinicians, but no alternative suggested.
Costelloe 2011 [121]	62	UK	P	6 months to 6 years	Parental estimates, parent vs. nurse measurement	< 2 kg	1	High	Findings: Parents can estimate weight accurately based on home measurements, but very few actually weigh children at home. Comments: Authors used doses of ibuprofen and paracetamol as endpoints: the applicability in emergency situations is uncertain. Limitations: Incomplete presentation of data. Very small sample size.
Geduld 2011 [122]	2832	South Africa	P	0 to 10 years	EPLS, Luscombe, BG, Broselow tape	< 10%*	1, 2	Low	Findings: Broselow tape and EPLS formula most accurate in this population. Careful titration of drugs and use of clinical judgement most important in using medications safely. Comments: Data from a poor community in South Africa. The accuracy of EPLS formula was the best ever reported while the accuracy of Broselow tape was on par with other reports. Only 4% of children excluded as too tall for the tape. The authors question whether the differences in accuracy of any weight estimation system are likely to affect outcomes. Limitations: Broselow tape not actually used and Broselow tape version not reported.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
Huybrechts 2011 [123]	275	Belgium	P	3 to 7 years	Parental estimates, parent vs. nurse measurement	None	1	Low	Findings: Parental estimates were most accurate when based on measurements made at home, rather than onguesses. Comments: Application to emergency weight estimations is uncertain as parental stress may negate this effect. Limitations: Targets were estimations of overweight or underweight. Incomplete presentation of data.
Kelly 2011 [124]	410	Australia	P	1 to 10 years	Luscombe, EPLS, Argall, BG	< 10%*	1, 2	Low	Findings: The Luscombe formula performed best of all the formulas. Comments: Study population BMI was 17. Same population as Nguyen 2007 and several other publications; all formulas performed poorly. Limitations: Incomplete presentation of data.
Knight 2011 [125]	657	USA	R	–	Broselow tape	None	1, 2	Low	Findings: Broselow tape performed poorly, potentially leading to under-resuscitation in all weight categories, especially in younger children. Drug doses correct in only about 50% of cases. Consensus opinion required whether to use IBW or TBW during resuscitation. Comments: Broselow tape 2007B. High incidence of obesity in study population. Limitations: Broselow tape not actually used. Incomplete presentation of data. No direct assessment of accuracy of weight estimation.
Luscombe 2011 [126]	64,197	UK	R	1 to 16 years	EPLS, Luscombe	None	1, 2	Low	Findings: The Luscombe formula outperformed the EPLS formula. Weight estimation is of paramount importance in resuscitation, therefore remembering one formula better than several. Comments: While the bias of the Luscombe formula was smaller, both formulas performed poorly. Limitations: Incomplete presentation of data. An inappropriate age range was used for formulas (up to 16 years).
Marlow 2011 [27]	140,314	UK	R	0 to 16 years	EPLS, Argall, L uscombe, BG	< 10%*	1, 2	Low	Findings: The EPLS formula was least accurate of commonly used formulas. The BG and Luscombe formulas were very similar and the best performers. No formulas showed acceptable accuracy, however. Comments: Abstract, with additional data supplied by author. This was a very large retrospective database study with good descriptive statistics. Limitations: Some incomplete data.
Rosenberg 2011 [38]	372	USA	P	0 to 14 years	healthcare provider guesses, Broselow tape	< 10%*	1, 2	Low	Findings: The Broselow tape was better than guesses by doctors, but not in obese children. Poor assessment of habitus by doctors. IBW suggested as the best target for estimation in obese kids. Comments: Broselow tape 2007B. 35% of study population obese or overweight. Mean BMI was 17.4. It is a reflection of how poorly the Broselow tape performed in obese children that doctor estimates were better; overall findings of Broselow tape accuracy similar to other studies. Limitations: Incomplete data presentation. Incorrect use of some statistics.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
Wells 2011 [127]	–	–	–	–	–	< 10%*	1	N/A	Findings: More than 50% of study population required habitus-modified weight estimation. Length-based systems should be used for weight estimation in emergencies. Comments: Editorial comment. Insufficient evidence exists to choose between IBW and TBW for drug dose calculations. Limitations: No reference standard for weight estimation included.
Abdel-Rahman 2012 [15]	1938	USA	R	0 to 16 years	EPLS, ARC, Argall, BG, Broselow tape, Cattermole, Leffler, Luscombe, Nelson, Shann, TJ, TK, Mercy method	None	1, 2	Low	Findings: The Mercy method performed significantly better than the age-based, length-based and MAC-based systems. Comments: This was the original development and validation study of the Mercy method. Good mix of overweight children, but few underweight kids in sample. Broselow tape unable to be used in more than one third of the sample (too tall for the tape). Limitations: Broselow tape not actually used and Broselow tape version not reported.
Ali 2012 [128]	1723	Trinidad	R	1 to 5 years	EPLS, Luscombe, new formula	< 10%†	1, 2	Low	Findings: All formulas performed similarly and all poorly, even the new formula derived from the study population. Comments: Only children aged 1 to 5 are included in study. Limitations: No validation sample for derived formula. Incomplete presentation of data.
Cantle 2012 [129]	2253	USA	R	1 to 10 years	EPLS, new APLS, CAWR-1, CAWR-2, MAC	None	1, 2	Low	Findings: CAWR-2 performed as well as other formulas. Age formulas performed less well in older children. MAC performed well in older children but poorly in younger children. Comments: Conference presentation. Study of CAWR in a Western population. No age formula actually performed well in any age group. Limitations: Some errors in presented data and calculations.
Elgie 2012 [130]	188	UK	P	10 weeks to 10 years	APLS, EPLS, Luscombe formulas based on clothing label age	None	1	High	Findings: Using the age on clothing label was more accurate than using actual age. Comments: This clever study used clothing size as a surrogate marker for habitus. The Luscombe formula performed slightly better than the EPLS formula. Limitations: Incomplete presentation of data.
Garwood 2012 [131]	1252	UK	R	1 to 16 years	EPLS, new formula	None	1	High	Findings: The Garwood formula was better than the EPLS formula. Comments: Abstract. Authors used assessment of mean bias only. Limitations: Limited statistical analysis made findings difficult to interpret. No validation of new formula.
Heyming 2012 [132]	491	USA	P	IQR 10 to 49 months	Broselow tape EMS vs. Broselow tape ED	None	1	High	Findings: The Broselow tape was an accurate tool to estimate weight. Comments: Broselow tape agreement between pre-hospital personnel and ED personnel only 70.1%. Limitations: Most of study population < 4 years. Incomplete presentation of ldata. Broselow tape version not reported.
Kelly 2012 [26]	–	–	–	–	–	None	1	N/A	Findings: Best systems to be used, in order of accuracy: parental estimates, Broselow tape, age-based formulas. Comments: Book chapter. Limitations: No mention of

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
									acceptable targets for weight estimation or weaknesses of age formulas. Subjective assessment of studies only.
Meguerdician 2012 [2]	–	–	–	–	–	None	1	N/A	Findings: The Broselow tape is the most consistent and reliable tool for weight estimation, but habitus-based weight adjustment for Broselow tape is logical. Parental estimates may rival the Broselow tape, but parents may be absent or uncertain, especially under stress. Underdosing might be prudent in emergencies. Comments: The author acknowledges that the degree of acceptable error is difficult to define. Limitations: No evidence provided of efficacy of Broselow tape in reducing medication errors. No suggestion of goal target accuracy required. Some controversial opinions about drug dosing.
Milne 2012 [33]	6361	Canada	R	IQR 1.6 to 7.7 years	Broselow tape	< 30%	1, 2	Low	Findings: Broselow tape effective, although significantly underestimated weight. Similar findings in urban and rural children. Comments: Broselow tape edition 2002A. The authors suggested that an error range of 30% is reasonable and that, ideally, technology should be developed to estimate weight. Limitations: Broselow tape not actually used—data taken from anthropometric measurements.
Park 2012 [28]	124,095	Korea	R	0 to 14 years	EPLS, Shann, Leffler, Nelson, BG, Broselow tape, Park	None	1, 2	Low	Findings: All methods underestimated weight, possibly because of a secular trend towards increasing BMI. The Park formula performed best and was accurate in children < 1 year of age. Comments: All methods assessed showed poor accuracy. Limitations: Incomplete presentation of data. Poor interpretation of bias vs. accuracy. Broselow tape was not actually used and Broselow tape version not reported.
Seddon 2012 [133]	599	UK	P	1 month to 12 years	EPLS, new APLS, Argall, BG, Luscombe	None	1, 2	Low	Findings: Mixed racial study population, no major racial differences found in performance of formulas. There was increased underestimation of weight with increasing age. The Luscombe formula performed better than EPLS and APLS in children 1 to 10 years of age; BG and APLS better outside these ages. Limitations: Limited statistical analysis and incomplete data—only measures of bias reported. Some data values were not credible.
Sinha 2012 [134]	118	USA	P	0 to 14 years	Broselow tape	None	1, 2	Low	Findings: It was possible to weigh children during trauma resuscitation. Comments: Broselow tape weight was compared to stretcher weight. Nearly 40% of children in the study could not be weighed—more often the sicker kids (not fully explained). There was no validation of the accuracy of weight measured during resuscitation. Poor performance by Broselow tape with 18% of children too tall for the tape. Limitations: Incomplete presentation of data. Broselow tape version not reported.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
Trakulsrichai 2012 [135]	595	Thailand	P	0 to 12 years	Broselow tape, parental estimates, growth charts	< 10%†	1, 2	Low	Findings: Family member estimation was most accurate, and the Broselow tape the most accurate of other weight estimation methods. Comments: Equal underestimation and overestimation by Broselow tape while family estimates tended to overestimate. Limitations: Incomplete presentation of data. Broselow tape version not reported.
Winship 2012 [42]	–	–	–	–	–	None	1	N/A	Findings: As average weight of children has increased so the accuracy of formulas has decreased; age-based methods will be unable to adjust to deal with future rises in average weights of children; no age-based weight formula accurate—might as well continue with EPLS formula. Comments: Semi-systematic review. Limitations: No suggestion of goal target accuracy required.
Wozniak 2012 [136]	777	Botswana	P	18 months to 12 years	EPLS, Luscombe, Theron, Cattermole, Broselow tape	< 10%†	1, 2	Low	Findings: Prediction models incorporating MAC and either tibia or ulna length performed extremely well. Age-based formulas were very inaccurate. Comments: Masters dissertation. Weight was markedly overestimated by formulas in this population with a high prevalence of HIV. Limitations: Incomplete statistics.
Abdel-Rahman 2013 [18]	624	USA	P	0 to 16 years	Mercy method two-dimensional, 3D, Broselow tape	None	1, 2	Low	Findings: Both two-dimensional and 3D Mercy methods outperformed the Broselow tape. Comments: Study population BMI was 17.3. Children much older than most weight-estimation studies included. Multicentre study—3 in the USA. One third of children were too tall for the tape. The 3D tape was significantly less accurate than the two-dimensional method. Limitations: Broselow tape not actually used and Broselow tape version not reported.
Abdel-Rahman 2013 [19]	976	USA	P	0 to 16 years	Mercy method, EPLS, Broselow tape, Luscombe, Nelson	< 20%	1, 2	Low	Findings: The Mercy method was accurate across a wider age range than other methods. Comments: Study population BMI was 17.6 with very few underweight children. Inter-rater assessment was generally reasonable, but two of the raters were inferior to the others. Limitations: Broselow tape not actually used and Broselow tape version not reported.
Akabarian 2013 [137]	403	Iran	P	0 to 14 years	Broselow tape, parental estimates	< 10%*	1, 2	Low	Findings: The Broselow tape was more accurate than parental estimates. Comments: Article in Arabic. Very good performance of Broselow tape compared to previous studies. Limitations: Exclusion criterion of weight > 35 kg limited the assessment of Broselow tape accuracy. Limited statistical analysis and data presentation. Broselow tape version not reported.
Cattermole 2013 [138]	171	Hong Kong	P	7 to 11 years	Broselow tape	None	1	High	Findings: Children "too tall for the tape" do not have full adult weight—this assumption would lead to an average 30% overestimation of weight in this study. About 40% of children were too tall for the tape. Comments: Selective evaluation of children who were too tall for Broselow tape. Interesting data which highlights a flaw in the Broselow tape methodology.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
									Broselow tape was not practically useful over the age of 10. An overestimation of 30% would be unacceptable; no mention of desirable target. Limitations: Relatively small sample size.
Graves 2013 [139]	37,114	USA	R	0 to 14 years	EPLS, new APLS, BG, Luscombe, Broselow tape	< 10%†	1, 2	Low	Findings: New APLS formula was best for infants, BG best for older children. Broselow tape wrong zone in up to 60% of children. Comments: Broselow tape 2007B. Despite criticisms of Broselow tape, it outperformed the formulas in every analysis. This study had some of the poorest performances of aged-based formulas in any study. A post hoc study population BMI was calculated to be 18 to 22. Limitations: Incomplete presentation of data. Broselow tape not actually used.
Hegazy 2013 [140]	508	Egypt	P	1 to 16 years	EPLS, Shann, Garwood formula	< 10%†	1, 2	Low	Findings: Garwood formula performed best, especially in older children. Comments: Sample population of cancer patients. Very poor performance of all formulas tested—none were close to acceptable accuracy. Limitations: Formulas used for children older than intended. Poor interpretation of findings.
House 2013 [141]	967	Kenya	P	0 to 14 years	Broselow tape, EPLS, Nelson	MPE < 10%	1, 2	Low	Findings: Broselow tape performed better than formulas and a measure of habitus assessment (e.g. MAC) was suggested. Broselow tape should be used rather than formulas. Comments: Broselow tape 2007B. Underestimation of weight predominated. Limitations: Incomplete statistical analysis and data presentation. Poor interpretation of results from previous studies; flawed outcome measures used (indicators of bias only).
Lim 2013 [39]	199	USA	R	0 to 17 years	healthcare provider guesses	< 20%†	1, 2	Low	Findings: The authors suggested that EMS personnel were generally accurate in estimating weights of children. Comments: The more severe the condition, the worse the weight estimation. Weight estimation appeared better than it actually was. Limitations: Incomplete presentation of data.
Loo 2013 [142]	875	Singapore	P	1 to 10 years	EPLS, Luscombe, Broselow tape	< 10%†	1, 2	Low	Findings: Broselow tape more accurate than formulas. EPLS underestimated, Luscombe overestimated weight, but EPLS performed better overall than Luscombe. Comments: Broselow tape 2007B. Mean BMI of study population was 15.9. Good statistics. Limitations: Broselow tape not actually used.
Suh 2013 [143]	105,072	Korea	R	–	Broselow tape 2005A, 2007B, 2011A	< 10%*	1, 2	Low	Findings: Broselow tape 2011A more accurate than older versions. Comments: Abstract. Large database sample. Limitations: Incomplete presentation of data. Broselow tape not actually used.
Wells 2013 [5]	453	South Africa	P	0 to 12 years	Broselow tape, PAWPER tape	< 10%†	1, 2	Low	Findings: PT performed better than Broselow tape in every category analysed and better than any previously published system. Comments: Broselow tape 2007B. Population with mixed under- and overweight. Multi-centre study of habitus-modified length-based weight estimation. Limitations: Assessment of body habitus based on visual estimation.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
Young 2013 [44]	207	USA	P	1 to 9 years	EPLS, Broselow tape, parental estimates, Luscombe, finger counting	< 10%*	1, 2	Low	Findings: Finger counting system as accurate as Broselow tape and more accurate than other formulas. Conceptually a simple system. To increase the accuracy weight-estimation systems may cause increased complexity and stress during resuscitations. Comments: The finger counting method is equivalent to formula Wt = 2.5 × age(years) + 7.5. Median BMI of study population was 17.2. Limitations: Incomplete presentation of data. Broselow tape not actually used and version not reported.
Abdel-Rahman 2014 [4]	–	–	–	–	–	None	1	N/A	Findings: Some experts have suggested that weight estimation cannot be accurate. This is likely to be to the disadvantage of children as dual length-and habitus-based can achieve acceptable accuracy. Comments: Brief review of weight estimation systems. Limitations: No target of weight estimation accuracy suggested; Mercy method recommended for environments where no scale available—no mention of emergency use.
Ackwerh 2014 [144]	10,488	USA	R	2 to 12 years	EPLS, Luscombe, Theron	< 10%*	1	High	Findings: Unreliable. Comments: Unclear statistical analysis. The reported results for age-based formulas were the best ever reported, but are completely incorrect. The graphically presented data contradict the other findings. Limitations: Incomplete presentation of data.
Allison 2014 [54]	2102	Australia	R	0 to 5 years	EPLS, BG, Luscombe, Argall, Nelson, Broselow tape, Sandell tape	None	1, 2	Low	Findings: Broselow tape was most accurate in this study. Comments: Broselow tape 2007 edition B. Aboriginal and Torres Strait Island children included (low weight-for-length). Only published study of Sandell tape. Limitations: Broselow tape not actually used. Incomplete presentation of data. Narrow age range evaluated. Very fat and very thin children excluded from study.
Batmanabane 2014 [49]	374	India	P	0 to 16 years	EPLS, ARC, Argall, BG, Broselow tape, Cattermole, Leffler, Luscombe, Nelson, Shann, TJ, TK, Mercy method	None	1, 2	Low	Findings: Mercy method performed well in Indian children, similar to that shown in Western populations. Comments: Good statistics. Overestimation of weight by all methods except Mercy. Limitations: Broselow tape not actually used and Broselow tape version not reported.
Chiengkriwate 2014 [145]	3869	Thailand	R	0 to 15 years	Broselow tape	< 10%*	1, 2	Low	Findings: Broselow tape underestimated weight in Thai children, more so in older children, similar to findings in Western populations. Comments: Broselow tape 2007 edition A. Broselow tape performance consistently at a PW10 of just below 60%. Limitations: Broselow tape not actually used.
Dicko 2014 [50]	473	Mali	P	0 to 16 years	Mercy, EPLS, ARC, Broselow tape, Nelson	None	1, 2	Low	Findings: Mercy method performed extremely well in this population in Mali, similar to its performance elsewhere in the world. Other methods overestimated weight. Comments: BMI of study population was 15.6, with 22% underweight and 1.7% overweight or obese. Good inter-rater reliability. Limitations: Broselow tape not actually used and version not reported.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
Eke 2014 [146]	370	Nigeria	P	1 to 12 years	APLS	None	1	High	Findings: The APLS formula underestimated weight in these Nigerian children. Comments: The findings are unreliable in view of the major methodological flaws. Limitations: Limited and incomplete data presentation.
Erker 2014 [147]	N/A	WHO/CDC reference centiles	R	1 to 12 years	EPLS, APLS, Luscombe, BG, Park, Shann, Nelson, novel habitus-modified formulas	None	1	High	Findings: Potentially enhanced accuracy of age-based equations by using a habitus-specific formula, if length-based methods not available. Comments: Three formulas for "thick", "normal" and "thin" children. Theoretical accuracy in derivation study. Limitations: No original data or validation of formulas. Incomplete presentation of data.
Flannigan 2014 [46]	10,081	UK	R	0 to 15 years	New APLS, Luscombe, novel formula	None	1, 2	Low	Findings: New APLS formula very inaccurate and should not be used. Weight should be adjusted according to a data table with the 5th and 95th weight-for-height values. Comments: Study performed in ICU patients. Limitations: Incomplete presentation of data. New regression formula untested, even with an internal validation sub-sample. Adjustment of weight estimation based on a "guess".
Omisanjo 2014 [148]	2754	Nigeria	P	1 month to 11 years	Best Guess, Nelson	MPE < ±5%	1, 2	Low	Findings: Neither formula was accurate in Nigerian children with a substantial overestimation of weight. Comments: This was one of the largest prospective studies of age-based formulas in the developing world. Limitations: Some limitations from incomplete statistics.
Abdel-Rahman 2015 [64]	400**	USA	P	2 month to 16 years	EPLS, Luscombe, DWEM, HCP guess, Broselow tape, 2DMT, 3DMT		1	High	Main results: "Real world" performance of weight estimation l ess accurate than when performed by experts. Calculation errors were common with age-based formulas. Weight and habitus were often underestimated by visual inspection. Usage errors with Broselow tape and Mercy tapes were common. Limitations: Not all modern weight estimation systems evaluated.
Asskaryar 2015 [57]	1185	India	P	1 month to 12 years	Broselow tape	< 10%[†]	1, 2	Low	Findings: Broselow tape significantly overestimated weight in Indian children. An 8% modification of the tape improved its accuracy. Comments: Broselow tape 2007 edition B. The improved performance was not substantially better than the original and was still below acceptable performance. Recalibrating bias alone is not enough when precision is low. Limitations: Incomplete presentation of data.
Badeli 2015 [149]	216	Iran	P	1 to 10 years	DWEM, Oakley, TJ, TK, MAC, Theron, Leffler, EPLS, healthcare provider guess, parental estimate	None	1	High	Findings: The authors reported that healthcare provider guesses and EPLS formula were more accurate than other methods. Comments: The findings are completely unreliable and the conclusions consequently unreasonable. Limitations: Incomplete presentation of data. Analysis of bias confused with accuracy.
Britnell 2015 [47]	376	New Zealand	P	5 to 10 years	EPLS, Shann, Theron, Broselow tape	< 10%[†]	1, 2	Low	Findings: Broselow tape more accurate than age-based formulas in children < 143 cm. Current acceptance of

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
									formulas needs to change. Large differences in accuracy of weight estimation in different ethnic groups. Comments: Broselow tape 2011 edition A. Large proportion of Pacific Island children in sample. Broselow tape could not be used in one fifth of children, but best ever performance of the Broselow tape in a study. Habitus was assessed but data not presented. Limitations: Incomplete presentation of data.
Chavez 2015 [13]	324	USA	P	1 month to 12 years	Broselow tape, PAWPER, APLS, MAC	None	1, 2	Low	Findings: Age-based formulas and the MAC formulas performed badly. The Broselow tape performed better and the PAWPER was most accurate overall, although not as accurate as in previous studies. Comments: Underestimation of obesity (and habitus score) caused underestimation of weight. High level of obesity in study population. Limitations: Broselow tape version not reported. Incomplete presentation of data. PW5 mistaken for PW10 by authors—confusing for readers.
Garcia 2015 [65]	1698	USA	P	–	PAWPER	None	1, 2	Low	Findings: The PAWPER did not perform as well as the original study. Comments: Most assignment of habitus score by nurses. Very high proportion of obesity. Limitations: Many incorrect assignments of habitus score. Children not fully undressed to assess habitus.
Khouli 2015 [150]	815	Mexico	P	0 to 12 years	Broselow tape	None	1, 2	Low	Findings: Broselow tape not accurate in Mexican population. Comments: Reasonable statistics. Both under- and overestimation found to be a problem. Nearly half of children had at least one colour zone error. Limitations: Broselow tape not actually used and Broselow tape version not reported.
Skrobo 2015 [151]	3155	Ireland	R	1 to 15 years	EPLS, Luscombe	< 15%	1, 2	Low	Main results: The Luscombe formula underestimated weight less than the EPLS formula. Comments: Increase in children's weight in "modern" populations related to increase in both lean body weight and adipose. Limitations: Only data on bias presented. Limited and incomplete data presentation.
Talib 2015 [152]	318	USA	P	0 to 18 years	APLS, Broselow tape, MAC formula, Mercy method	None	1	High	Main results: Mercy method performed better than Broselow tape and age-based formulas in children with Down syndrome. Comments: No system performed well in this population. Limitations: Limited and incomplete data presentation.
Young 2015 [153]	207	Philippines	P	1 to 9 years	EPLS, APLS, Luscombe, BG, finger counting, Broselow tape	None	1	High	Findings: Broselow tape performed best in this population, updated APLS formula performed worst. Comments: Broselow tape 2011 edition A. Limitations: Broselow tape not actually used. Incomplete presentation of data.
AlHarbi 2016 [154]	3537	Saudi Arabia	P	1 month to 12 years	Broselow tape 2007B and Broselow tape 2011A	None	1	High	Findings: Broselow tape 2011A performed better than Broselow tape 2007B in this population. The authors suggested that the tapes were accurate. Comments: The method of statistical analysis does not support the conclusions drawn in this paper. Limitations: Unclear if the Broselow tape was used or

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
									if derived from length measurements. Limited and incomplete data presentation.
Aliyu 2016 [155]	300	Nigeria	P	0 to 5 years	Broselow tape, APLS	None	1, 2	Low	Findings: Broselow tape and APLS formulas performed well in this population. Comments: Contrary to the authors' conclusions, although the Broselow tape and APLS formula performed similarly, they were both inaccurate. Limitations: Not clear if Broselow tape used or if derived from length measurements. Broselow tape edition not reported. Limited and incomplete data presentation.
Britnell 2016 [51]	376	New Zealand	P	5 to 10 years	Novel weight lookup table	< 10%	1	Low	Main results: Length- and habitus-based model better than age- and habitus-based model. Model outcome similar to Broselow tape. Comments: Useful additional information on value of habitus-based methods of adjusting weight estimations. Limitations: Limited and incomplete data presentation.
Clark 2016 [62]	583	Sudan	R	6 months to 5 years	Broselow tape	< 10%**	1,2	Low	Findings: The Broselow tape performed very poorly. Comments: Abstract. Study in South Sudan, the "hungriest place on earth" where 61% of study population was malnourished. There was up to a two colour-zone overestimation in severely malnourished children with only 26% agreement in normally nourished children. Very poor performance of the Broselow tape. Dangerous overestimation of weight in undernourished children. Limitations: Broselow tape not actually used and Broselow tape version not reported.
Carasco 2016 [156]	–	–	–	–	–	None	1	N/A	Main results: EPLS formula may be better suited to identifying ideal body weight rather than total body weight. Comments: Systematic review. Ideal body weight should only be used in obese children. Limitations: No weight estimation data.
Chassee 2016 [157]	197	USA	R	0 to 12 years	911 caller weight estimation	< 20%	1	Low	Main results: Family members could frequently provide a weight estimate to the 911 operator. Comments: More than 20% of cases had a weight estimation error > 20%. Limitations: Limited and incomplete data presentation.
Georgoulas 2016 [21]	300	South Africa	P	1 month to 12 years	Broselow tape, PAWPER, Wozniak, Mercy	< 10%*	1, 2	Low	Findings: PAWPER tape performed best, but good performances from Wozniak and Mercy methods. Broselow tape was worst performer. Comments: Unpublished data. Broselow tape 2011 edition A. Poor population with high proportion of underweight children. First comparative study of these methodologies. Relatively weaker performance by all methods in infants, but Wozniak especially was very weak. Limitations: Assessment of body habitus done by single researcher.
Jung 2016 [16]	906	Korea	P	0 to 17 years	Broselow tape, novel device	None	1, 2	Low	Main results: The novel device performed better than the Broselow tape in all outcome measures and was quicker to use. Comments: Both devices underestimated weights. Overweight and underweight children were often misclassified

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
									into wrong habitus category. Broselow tape 2011 edition A. Limitations: Device is not commercially available. Incomplete statistical analysis
Jung 2016 [158]	–	–	–	–	–	None	1	N/A	Non-systematic review. Main results: Weight frequently underestimated in older children. Differences in body habitus account for this error, which needs further research. Comments: Article in Korean. Limitations: Non-systematic review. No quantitative assessment of methodologies.
Lowe 2016 [55]	3018	USA	R	0 to 13 years	Broselow tape, Handtevy tape	None	1, 2	Low	Main results: The Broselow tape performed better than the Handtevy tape in all outcome measures. Both tapes underestimated weights. Comments: Although the authors recognised the increased accuracy when obese children were excluded, they still erroneously advocated "recalibration" of the tapes. Broselow tape 2011 edition A. Limitations: Tapes not actually used. Limited and incomplete data presentation.
Mishra 2016 [159]	603	India	P	0 to 10 years	Broselow tape	None	1	Low	Findings: Broselow tape performed best in smallest children. Comments: Broselow tape 2007 edition B. Only colour zone accuracy was assessed, which was poor in children > 18 kg. Limitations: Limited and incomplete data presentation.
Nosaka 2016 [160]	237	Japan	P	0 to 10 years	EPLS, Park, Broselow tape, parental estimates	None	1, 2	High	Main results: Parental estimates were most accurate, then Broselow tape then APLS formula. Mothers' weights, as well as Broselow tape and EPLS estimations were extremely accurate. Comments: The BMI for this population suggests that very few obese children were included and the sample was skewed towards very young children. Broselow tape 2007 edition B. Limitations: Limited and incomplete data presentation.
Ralston 2016 [29]	453,990	Multicentre International	R	6 months to 5 years	Broselow tape, MAC, height + MAC model	None	1, 2	Low	Findings: A novel model incorporating height and MAC was the most accurate. Broselow tape 2011 edition A less accurate than 2007 edition B in this population. Comments: Broselow tape 2007 edition B and 2011 edition A. A very large multinational database study. Limitations: Broselow tape not actually used.
Sahar 2016 [161]	1163	Malaysia	P	0 to 12 years	Broselow tape	None	1, 2	Low	Findings: Broselow tape underestimated weight in small children and overestimated weight in older children. It was not accurate. Comments: As with studies elsewhere there was a large variation in accuracy. Limitations: Broselow tape version not reported. Limited and incomplete data presentation.
So 2016 [162]	4600	Hong Kong	R	1 to 12 years	APLS, EPLS, BG, Luscombe, CAWR-1, CAWR-3	None	1, 2	Low	Main results: A new CAWR formula was more accurate than other formulas in this study. Comments: No formula was very accurate and the differences between formulas were generally small. Limitations: Data very skewed towards younger children. No inclusion of length-based methods.

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
So 2016 [163]	4178	Hong Kong	R	1 to 9 years	Finger counting	None	1, 2	Low	Main results: The finger-counting formula was more accurate than other formulas. Comments: Same data was used for So 2016, but with a narrower age-restriction and the new finger-counting formula. Limitations: Duplicate data. Only the data on finger counting was added.
Tanner 2016 [17]	178	USA	P	2 to 18 years	Broselow tape	None	1	High	Main results: The Broselow tape was very inaccurate in overweight and obese children, who made up more than half the sample. Habitus assessment was poor by parents and nurses, but better by the principal investigator. Comments: Using a corrective formula based on waist circumference to adjust the Broselow tape weight improved accuracy in obese children. Limitations: Broselow tape version not reported. Limited and incomplete statistical analysis. Only included children who fell within the length limitations of the tape.
Young 2016 [9]	–	–	–	–	–	–	1	N/A	Systematic review. Main results: Parental estimates were the most accurate technique followed by length- and habitus-based methods. Broselow tape more accurate than age formulas. Limitations: No quantitative data analysis.
Bowen 2017 [164]	1381	Zambia	P	0 to 14 years	Broselow tape, APLS, EPLS, ARC, Argall, BG, CAWR, Garwood, Leffler, Luscombe, Michigan, Nelson, Park, Shann, Theron	< 10%*	1, 2	Low	Findings: Broselow tape performed better than every formula in this population, BG and Michigan formulas performed worst. Comments: None of the methods were accurate, and all methods overestimated weight. Limitations: Broselow tape not actually used. Broselow tape version not reported.
O'Leary 2017 [22]	199	Australia	P	0 to 14 years	PAWPER, APLS, Luscombe, BG, Broselow tape, Mercy	< 10%*	1, 2	Low	Findings: Age-based formulas performed badly. The Broselow tape and Mercy method performed significantly better and the PAWPER was most accurate overall. Comments: All systems performed worst in infants. Reasonably good performance of Broselow tape, possibly because of newer edition used. Limitations: Broselow tape version not reported.
Reilly 2017 [165]						None	1	N/A	Commentary. Main results: MAC is useful as an adjunct to estimate weight. Limitations: No original data.
Samerchua 2017 [68]	430	Thailand	P	0 to 12 years	Parental estimations, Mercy method, Broselow tape, EPLS, APLS	None	1, 2	Low	Main results: Parental estimates were most accurate, followed by the Mercy method. Comments: Parents could only provide an estimate in 80% of cases—the Mercy method was the only method that could be used in all children. Limitations: It was not clear what training the raters had received.
Trainarongsakul 2017 [166]	345	Thailand	P	0 to 15 years	RAMA Ped card	None	1	Low	Main results: RAMA Ped card had a fair correlation with weight. Comments: The use of only correlation data made the findings hard to interpret. Limitations: Limited and incomplete data presentation.
Waseem 2017 [167]	538	USA	R	0 to 8 years	Broselow tape	None	1	N/A	Main results: Broselow tape underestimated weight in older children. It was not accurate in nearly half the population. Broselow tape 2011 edition A. Comments: Many underweight

Table 1 Studies included in the qualitative review and quantitative meta-analysis *(Continued)*

Author and date	Study size (*N*)	Country	Design	Patient ages	Estimation techniques evaluated	Target	Arm	Risk of bias	Major findings; comments; major limitations
									and obese children included which cause the poor performance. Limitations: Broselow tape not actually used. Limited and incomplete data presentation. Only included children who fell within the length limitations of the tape.
Wells 2017 [20]	328	South Africa	P	0 to 16 years	Broselow tape, PAWPER, Wozniak, Mercy	< 10%†	1, 2	Low	Findings: The Broselow tape performed poorly in this study, the Wozniak method and the Mercy method showed intermediate accuracy and the PAWPER was most accurate overall. Comments: This was a population with many older children and children with deviations from "average" weight-for-length. The PAWPER XL tape worked well in this population. Limitations: The Mercy method was used in a simulated resuscitation setting (supine children), which may have affected its accuracy.
Wells 2017 [11]	1085	South Africa	P	0 to 16 years	20 age-based formulas, Traub-Johnson and Traub-Kichen formulas including habitus-modified age- and length-based formulas	< 10%	1, 2	Low	Findings: No age- or length-based formula performed acceptably well. Age-based habitus-modified formulas also performed poorly, but length-based habitus modified formulas were, surprisingly, extremely accurate. Comments: Length-based formulas can predict both TBW and IBW, but are mathematically complex. Limitations: Habitus assessment by visual estimation is subjective.
Wells 2017 [31]	13,134	USA	R	0 to 18 years	PAWPER XL-MAC method	< 10%	1, 2	Low	Findings: The PAWPER XL-MAC method was very accurate in data from both the USA and South Africa. Comments: This sytem was completely objective with no visual assessment of habitus. Limitations: Retrospective study
Whitfield 2017 [14]	–	–	–	–	–	–	–	–	As for Wozniak 2012

A brief summary of findings as well as a short commentary on significant aspects is included. In the description of target accuracy, some studies used an implied target (indicated by an asterisk (*)) and some expressed a clear, strong opinion (indicated by a dagger (†)). Whether studies were entered into arm 1 (qualitative arm) only, or both arm 1 and arm 2 (meta-analysis) is indicated. The final assessment of the risk of bias is also indicated. Abbreviations: *APLS* Advanced Paediatric Life Support formula, *ARC* Australian Resuscitation Council formula, *BG* Best Guess formula, *BT* Broselow tape, *CAWR* Chinese age-weight rule formula, *DWEM* devised weight-estimating method, *EPLS* European Paediatric Life Support formula, *TJ* Traub-Johnson formula, *TK* Traub-Kichen formula

Table 2 Summary and description of weight estimation methodologies described in the literature

	Name	Formula	Restrictions/limitations/acceptable accuracy benefits	
Age-based and length-based formulas	Ali formula	$Wt = (2.5 \times Z) + 8$	Derived in a Trinidadian population of children ≤ 5 years of age in 2012. No validation studies to date. Age restriction 1 to 5 years of age.	
	Argall formula	$Wt = (3 \times Z) + 6$ *or* $[Wt = 3 \times (Z + 2)]$	Developed from a small UK study in 2003 (300 children). Generally found to underestimate weight, more so in older and heavier children. Age restriction 1 to 10 years of age.	
	Advanced Paediatric Life Support formula (APLS) (new)	$Wt = \frac{z}{2} + 4$	For infants ≤ 12 months of age	Derived in a UK population and adopted in 2011 by the Advanced Life Support Group from a combination of the original APLS and the Luscombe formulas. It was untested and unvalidated at the time of adoption. Generally overestimates weight. Age restriction birth to 12 years of age.
		$Wt = (2 \times Z) + 8$ *or* $[Wt = 2 \times (Z + 4)]$	For children aged 1 to 5 years	
		$Wt = (3 \times Z) + 7$	For children aged 6 to 12 years	
	Australian Resuscitation Council formula (ARC)	$Wt = 3.5$	At birth	Adopted by the ARC in Australia in 1996. Same as New Zealand Resuscitation Council formula. Generally underestimates weight, more so in older and heavier children. Differing accuracy in different ethnic, socio-economic and international populations. No specific age restriction noted.
		$Wt = (2 \times Z) + 8$	For children aged 1 to 9 years	
		$Wt = 3.3 \times Z$	For children 10 years and over	
	Best Guess formulas (BG)	$Wt = \frac{Z+9}{2}$	For infants ≤ 12 months of age	Also known as the Tinning formulas. Derived in Australian population in 2007 from a retrospective database study of more than 70,000 children. Generally overestimates weight, especially in poorer populations. Has been evaluated in several validation studies with mixed results.
		$Wt = (2 \times Z) + 10$ *or* $[Wt = 2 \times (Z + 5)]$	For children aged 1 to 5 years	
		$Wt = 4 \times Z$	For children aged 6 to 14 years	
	Bicer formula	$Wt = (3 \times Z) + 6$ *or* $[Wt = 3 \times (Z + 2)]$	For children aged 3 to 6 years	Although these formulas are mentioned and evaluated in the Bicer study, the analysis and reporting is fatally flawed and cannot be evaluated; the origin of the first formula of the set is the same as the Argall formula. Age restriction proposed by Bicer to be 3 to 18 years.
		$Wt = (4 \times Z) - 3$	For children aged 7 to 18 years	
	Chinese age-weight rule 1 (CAWR-1)	$Wt = (3 \times Z) + 5$	For children aged 1 to 10 years	Developed in Hong Kong for ethnic Chinese children in 2011 from a sample of 1248 children. Age restriction 1 to 10 years (although developers advise use with caution over 7 years).
	Chinese age-weight rule 2 (CAWR-2)	$Wt = \frac{(Z \times 7) + 25}{3}$	For children aged 1 to 6 years	
		$Wt = (4 \times Z) - 4$	For children aged 7 to 10 years	
	European Paediatric Life Support formula (old APLS formula) (EPLS)	$Wt = 2 \times (Z + 4)$ *or* $[Wt = (2 \times Z) + 8]$	Original population and date of derivation unclear. Generally underestimates weight, more so in older and heavier children. Differing accuracy in different ethnic, socio-economic and international populations. Age restriction 1 to 10 years of age	
	Garwood formula	$Wt = \frac{z}{4} + 6$	Developed in a UK population from a sample of 1252 children in 2012. The initial validation study was fatally flawed, but this formula has been subjected to a validation study subsequently (showing poor performance). For children aged 1 to 16 years.	
	Leffler formulas	$Wt = \frac{Z+8}{2}$	For children <1 year of age	Also known as the Tintinalli formula, the original origin is unclear, but became popular after the Leffler study in 1997. Overestimates weight in younger children (≤ 6 years) and underestimates weight in older children (> 6 years).
		$Wt = (2 \times Z) + 10$	For children aged 1 to 10 years	

Table 2 Summary and description of weight estimation methodologies described in the literature *(Continued)*

	Name	Formula	Restrictions/limitations/acceptable accuracy benefits	
	Luscombe formula	$Wt = (3 \times Z) + 7$	Developed in the UK in 2007 from a large database of nearly 14,000 children. Underestimates weight in most populations studied, but significantly overestimates weight in populations from developing countries. Age restriction 1 to 10 years.	
	Michigan formula	$Wt = (3 \times Z) + 10$	Derived by Ackwerh in 2010, but has not been evaluated fully.	
	Nelson formulas (originally Weech's formulas)	$Wt = \frac{Z+9}{2}$	For infants 3 to 12 months	As described in Nelson's Textbook of Paediatrics. The origin is probably from Weech's formulas, first reported in 1954 in the USA. The Weech formula is still in use today as one of the standard measurement denominators for determining underweight status. Weight most often overestimated in infants and older children (> 6 years) and underestimated in younger children (≤ 6 years).
		$Wt = 2 \times (Z + 4)$	For children aged 1 to 6 years	
		$Wt = \frac{(Z \times 7) - 5}{2}$	For children aged 7 to 12 years	
	Park formulas	$Wt = \frac{Z+9}{2}$	For infants ≤12 months of age	Derived in Korean population from a large database study (nearly 125,000 children). Poor accuracy in older children (> 6 years).
		$Wt = (2 \times Z) + 9$	For children aged 1 to 4 years	
		$Wt = (4 \times Z) - 1$	For children aged 5 to 14 years	
	Shann formulas	$Wt = (2 \times Z) + 9$	For children aged 1 to 9 years	Used in Australasia primarily. Origin is unclear. Underestimates weight increasingly with increasing age.
		$Wt = (3 \times Z)$	For children aged >9 years	
	Theron formula	$Wt = e^{(0.175571 \times Z) + 2.197099}$	Derived in 2005 in New Zealand from a small study of 900 children that included a large number of Pacific Island children (high weight-for-age). The developers intended it for use in children high in the weight-for-age centiles. Age restriction 1 to 10 years. Overestimates weight in most populations.	
	Unknown	$Wt = (3 \times Z) + 8$	This formula is mentioned in some weight-estimation studies with no reference to its origin. It is not known what restrictions apply. Mentioned in Dearlove, Bicer.	
	Traub-Johnson formula (TJ)	$Wt = 2.05 \times e^{0.02X}$	Derived in 1980 from USA national growth data from 1959. This formula was used to estimate ideal body weight and adjusted body weight, which were used interchangeably. The formula was intended to estimate the 50th centile of weight-for-height. Underestimates total body weight. For children aged 1 to 18 years.	
	Traub-Kichen formula (TK)	$Wt = 2.396 \times 1.0188^{X}$	Derived in 1983 in the USA from data from more than 20,000 children in the National Centre for Health Statistics database. The formula was intended to estimate the 50th centile of weight-for-height which the developers regarded as an approximation of ideal body weight. Underestimates total body weight. For children over 74 cm and aged 1 to 17 years.	
Other length-based systems	Broselow tape (BT)	Weight estimated directly by placing tape next to child and measuring from head to heel. The estimated weight and colour zone is read off the tape.	Developed in 1985 in the USA from US growth data and first validated in a sample of just over 900 children in 1988. Several changes have been made over the years: the latest version is the 2011A edition. Underestimates weight except in populations with a high prevalence of poor nutrition. Inaccuracy increases with increasing length/weight. Increased underestimation of weight in obese and overweight individuals. Substantial number of children "too tall for the tape" but who are not at adult weight. Length restriction 46 to 143 cm. Maximum weight estimation 36 kg.	

Table 2 Summary and description of weight estimation methodologies described in the literature *(Continued)*

	Name	Formula	Restrictions/limitations/acceptable accuracy benefits
	Blantyre tape	Weight estimated directly by placing tape next to child and measuring from head to heel. The estimated weight is read off the tape.	Developed in Malawi using values 85% of the 50th centile of the American National Centre for Health Statistics weight-for-length growth charts. Validated on a sample of 729 children. The developers reported a reasonable accuracy between 4 and 16 kg, but the reporting of data was fatally flawed and is unverifiable. Length restriction of 45 to 130 cm.
	Oakley table	Age or length is used to estimate weight from a graph.	Developed in the USA in 1988 from averaged boy-girl medians of unspecified growth charts. Overestimates weight in infants and older children (> 6 years). Age restriction 0 to 14 years and l ength restriction 50 to 160 cm.
Habitus-modified systems	Erker formulas	$Wt = (2 \times Z) + 6$ $Wt = (3 \times Z) + 6$ $Wt = (4 \times Z) + 6$	For "thin" children For "normal" (average) children For "thick" (fat) children Developed in 2014 in the USA using regression formulas to estimate the 5th, 50th and 95th centiles of the Centre for Disease Control weight-for-age growth charts. Has not yet been shown to be accurate.
	Yamamoto formulas	*See reference 11 different logarithmic formulas.*	Developed in 2009 in the USA from a sample of 542 children. A different length-based formula is selected for one of five (under 3 years of age) or six (over 3 years of age) icons, which represent an assessment of the body habitus. The reporting of the validation against the Broselow tape is flawed and does not permit verification. This technique has not been subsequently validated.
	Wozniak formulas	$Wt = (1.443 \times U) + (1.596 \times M) - 32.963$ $Wt = (0.86 \times T) + (1.715 \times M) - 30.426$	Developed in Botswana in 2012 from a sample of 777 children with a high prevalence of HIV infection and growth retardation. Measurements of mid-arm circumference and ulna length or tibia length are used to estimate weight using the formula. The accuracy of the method decreases in children <10 kg and children > 40 kg.
	Devised weight estimating method (DWEM)	Length measured and then habitus assessed as "Slim", "Average" or "Heavy" and weight read off a chart. A pre-marked tape was devel oped but is not widely available.	Developed in 1986 in the USA based on standard growth data available in 1983 and validated in a small sample of 258 children. Underestimates weight, especially in taller children. Length restriction 50 to 175 cm. Maximum weight estimation 70 kg.
	PAWPER tape	Weight estimated directly by placing tape next to child and measuring from head to heel. A habitus score (1 to 5) is assigned to the child based on body habitus (1 = very thin, 3 = average, 5 = very fat). The estimated weight for that length and habitus score is read off the tape.	Developed in 2004 in South Africa based on WHO weight-for-length growth charts and validated on a sample of 453 children in 2013. Estimates weight uniformly across length range of tape. Performs well in children who are under- or overweight. Length restriction 43 to 153 cm. Maximum weight estimation 47 kg. The extended PAWPER tape accommodates children up to 180 cm in length, a maximum weight estimation of 116 kg and with a 7-point habitus score assessment (habitus scores 6 and 7 were added to accommodate children above the 95% centile of weight-for-length, i.e. for obese and severely obese children).
	Mercy method (MM)	Humerus length and mid-arm circumference are measured and then used to determine "segmental weights" from a table. Specifically desgined tapes "2D" and "3D" tapes may be used which eliminates the need for a data table.	Developed in the USA from a database of 19,625 children and validated across several centres in 2012, 2013 and 2014, including in developing countries. Consistently good weight estimation across age and habitus ranges. Decreased accuracy in younger children (< 2 years).

Table 2 Summary and description of weight estimation methodologies described in the literature *(Continued)*

	Name	Formula	Restrictions/limitations/acceptable accuracy benefits
Other	Cattermole MAC formula	$Wt = (M - 10) \times 3$	A mid-arm circumference-based formula developed in Hong Kong ethnic Chinese children in 2010 from a sample of 1370 children. Decreased accuracy in children < 6 years and underestimation of weight in older children. Recommended by developers to restrict use to children aged 6 to 11 years.
	Haftel formula	$Wt = (5.176 \times LW) + 3.487$	The "hanging-leg weight" formula developed in the USA in 1990 from a small sample of 100 anaesthetised children aged 2 months to 15 years. The accuracy of weight estimation was worst in infants and increased with age. No subsequent studies have been reported.
	Bavdekar formula	$Wt = -5.15 + (FL \times 1.35)$	Developed in India in 2006 from a sample of 500 infants < 2 years of age. Fatal flaws in the methodology do not permit the interpretation of the accuracy of this formula. No subsequent studies have been reported. For infants ≤ 2 years.

Methods not shown include the Carroll method, the Sandell and Handtevy tapes (insufficient data) and neonatal weight estimation applications (out of scope). Only weight estimation systems that had more than one article assessing their functioning were considered for inclusion into the meta-analysis. Abbreviations: *Z* age in years, *z* age in months, *X* height or length in centimetre, *M* mid-arm circumference in centimetre, *LW* hanging leg-weight in kilogramme, *FL* foot length in centimetre, *U* ulna length in centimetre, *T* tibial length in centimetre

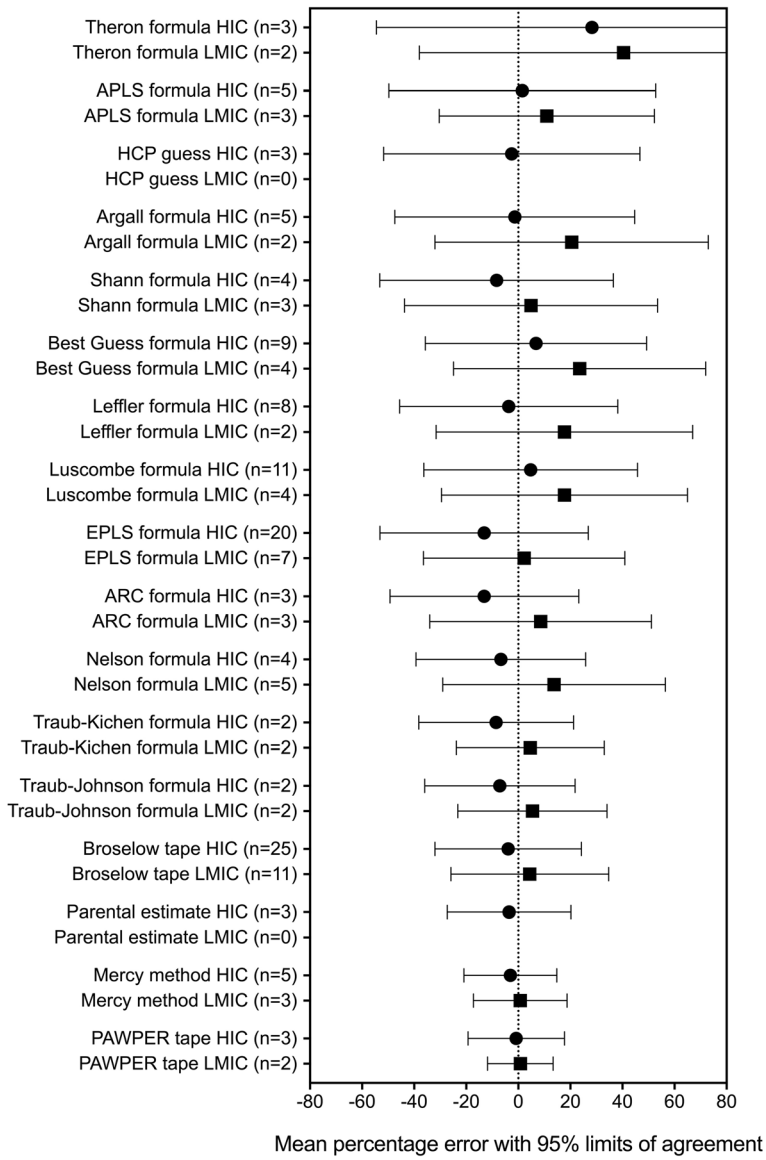


Fig. 2 Forest plot showing the bias and precision data of the major weight estimation systems evaluated

two-dimensional systems were the most accurate. Despite the difference in bias between high-income countries and low- and middle-income countries for the one-dimensional systems, the overall accuracy was similarly poor. If a PW10 of 70% were used as a benchmark of acceptable accuracy, only the PAWPER tape and the Mercy method would have achieved acceptable accuracy, with parental estimates close behind. When examining the PW20 data in Table 3, only the PAWPER tape (96.6%) and the Mercy method (95.3%) met the acceptability criteria suggested by Stewart of a PW20 > 95% [30]. The PW20s for the Broselow tape, parental estimates and a value calculated for pooled age-based formulas were 81.2, 87.1 and 65.0%, respectively.

The results of the subgroup analyses are shown in Table 4.

Figure 4 shows the results of direct statistical comparisons between weight estimation systems from studies where paired data could be pooled, using non-parametric measures of accuracy (PW10 data). The full analyses are available in Additional file 2: Figure S1. There was little difference between the accuracy of the different age formulas. Length-based methods were always more accurate than age-based methods, and two-dimensional methods were more accurate than one-dimensional methods. On direct comparison, but with data from only two studies, the PAWPER tape was significantly more accurate than the Mercy method.

Parental estimates were significantly more accurate than the Broselow tape, but there was no data for direct comparison with any two-dimensional system.

Discussion

Summary

The quality of the evidence from the contributing studies was generally good, and the number of studies that could be included allowed for a comprehensive analysis of the data. The underlying risks of bias, while present, were considered not sufficient to alter the overall findings. Additional information on parental estimations of weight in different populations and circumstances is also required, as well as a comparison with the two-dimensional weight estimation systems.

The implications of the results for clinical practice and future research are profound: age-based formulas, along with healthcare provider guesses, were the least accurate of all weight estimation systems. They should not be used or taught. Similarly, one-dimensional length-based systems, while widely used and advocated by advanced life support organisations, were simply not accurate enough. The future challenges will be to develop two-dimensional systems, which produced the most accurate weight estimations, to be safe, quick and easy-to-use during emergency care.

Many articles on weight estimation have been—and continue to be—published without any clear indication if the results achieved, and the weight estimation systems tested, were actually good or bad. This meta-analysis has provided some useful findings which could guide researchers and decision-makers on which systems to use in clinical practice and which to explore in further research. It has also provided some perspective on the performance of weight estimation systems in high-income and low-and middle-income populations, which is important as most weight estimation systems have been developed in high-income countries and have the potential to be dangerous if used inappropriately.

A benchmark for weight estimation systems

What degree of under- or overestimation of weight is dangerous to a child when calculating drug doses is not known [34, 35]. Many of the drugs used in paediatric emergencies have not been adequately studied to determine optimal dosing ranges. Moreover, the consequences of overestimating or underestimating weight (and therefore dose) will differ between different drugs, different patients and different clinical scenarios [36–39]. The final dose will be strongly influenced by the clinical situation and the discretion of the treating doctor, but an accurate and reliable weight estimation would still be required to provide the starting point to allow for dose modifications. Some authors regard the need for a highly accurate weight estimation as debatable. Other argue that any factors potentially impacting on patient safety must be addressed and minimised, especially in the light of compounded errors in drug dose calculations [40].

In the qualitative arm of the systematic review, we found no objective evidence to support any particular target or system by which to assess the adequacy of weight estimation methodologies. The failure to define outcome measures on how accurately a weight estimation method must perform is methodologically unsound, however. This is important as the use of a system known to be inaccurate, or inferior to another system is not good medical practice [1]. There are clearly factors other than accuracy to consider when selecting the most appropriate weight system to adopt including the complexity and cognitive load generated by the system, the vulnerability to human factor errors and its ability to interface with a drug dosing guide [41]. This needs further research.

Despite the lack of objective evidence, some reference standard is still required. A large number of articles implied or stated explicitly that an individual estimation of weight within 10% of actual weight is desirable, but only three articles provided a benchmark by which to judge a weight estimation system. The suggested criteria were that, to be considered accurate, 70% of weight estimates must be within 10% of actual weight and 95% of weight estimates must fall within 20% of actual weight [11, 30, 31]. Since the newest two-dimensional systems have shown the capability to repeatedly achieve this standard, it could, therefore, be considered a reasonable benchmark to propose to assess the adequacy of weight estimation systems in the future.

Meta-analysis data: the accuracy of weight estimation systems

Age-based weight estimation

The age-based formulas were the least accurate and worst-performers of all the weight estimation methods. There are multiple reasons for the inaccuracy of age-based formulas: a large biological variability in weight-for-age; a non-linear relationship between weight and age; and differences between populations with different ethnic groups and different levels of nutrition [10]. We found that age-based formulas have never been shown to perform better than length-based systems. Despite this, many authors still regard the EPLS formula as the “gold standard” for weight estimation and age-formulas are still taught on advanced life support courses [42, 43]. Some authors also still support the use of age-based formulas because of their ostensible simplicity, because they require no equipment to function and they allow advanced preparation if emergency services

Table 3 Weight estimation meta-analysis summary data, showing both fixed effects (FE) and random effects (RE) data

		Random effects outcomes									Fixed effects outcomes								
		MPE	LLOA	ULOA	Studies	*N*	PW10	PW20	Studies	*N*	MPE	LLOA	ULOA	Studies	*N*	PW10	PW20	Studies	*N*
APLS (new)	All	5.1	−42.5	52.6	8	19,029	33.6	56.1	13	49,709	6.1	−44.1	56.2	8	19,029	28.3	58.8	13	49,709
APLS (new)	HIC	1.5	−49.7	52.8	5	17,318	33.7	56.1	9	46,709	5.8	−44.7	56.3	5	17,318	28.0	58.8	9	46,709
APLS (new)	LMIC	10.9	−30.4	52.3	3	1711	33.4	50.6	4	3000	9.1	−37.2	55.4	3	1711	33.8	56.9	4	3000
ARC	All	−2.3	−41.6	37.1	6	6334	35.9	64.2	7	7405	−8.3	−50.5	33.8	6	6334	35.9	65.3	7	7405
ARC	HIC	−13.1	−49.3	23.2	3	4572	35.1	66.7	3	4572	−13.4	−50.1	23.2	3	4572	34.3	64.9	3	4572
ARC	LMIC	8.6	−34.0	51.1	3	1762	36.5	62.4	4	2833	4.9	−39.4	49.2	3	1762	38.6	65.9	4	2833
Argall formula	All	4.9	−43.0	52.8	7	145,823	31.5	51.1	10	147,341	−4.7	−47.3	37.9	7	145,823	38.7	63.1	10	147,341
Argall formula	HIC	−1.3	−47.4	44.7	5	144,691	34.7	54.3	7	145,257	−4.8	−47.2	37.6	5	144,691	38.9	63.2	7	145,257
Argall formula	LMIC	20.5	−32.1	73.1	2	1132	24.1	46.8	3	2084	14.3	−37.4	66.1	2	1132	28.4	54.1	3	2084
Best Guess formula	All	12.0	−32.4	56.3	13	284,798	33.8	58.2	18	324,035	5.8	−35.2	46.8	13	284,798	39.2	65.1	18	324,035
Best Guess formula	HIC	6.8	−35.7	49.3	9	277,790	37.7	63.5	14	318,478	5.5	−35.2	46.1	9	277,790	39.5	65.3	14	318,478
Best Guess formula	LMIC	23.6	−24.9	72.1	4	7008	20.2	46.4	4	5557	19.3	−25.5	64.1	4	7008	21.9	58.5	4	5557
CAWR-1	All	−3.3	−45.4	38.9	4	9064	34.6	65.7	5	10,016	−5.2	−48.2	37.7	4	9064	34.4	65.2	5	10,016
CAWR-1	HIC	−5.6	−47.1	35.9	3	8101	34.1	68.1	3	8101	−6.3	−48.6	36.0	3	8101	34.2	65.9	3	8101
CAWR-1	LMIC	3.9	−40.4	48.2	1	963	35.5	62.0	2	1915	3.9	−40.4	48.2	1	963	35.5	62.0	2	1915
CAWR-2	All	1.2	−38.6	40.9	3	4384	41.4	76.2	3	4384	−1.0	−42.1	40.0	3	4384	42.2	76.3	3	4384
CAWR-2	HIC	−2.2	−41.2	36.8	2	3501	41.7	79.0	2	3501	−3.3	−43.1	36.5	2	3501	42.6	77.7	2	3501
CAWR-2	LMIC	7.9	−33.3	49.1	1	883	40.7	70.8	1	883	7.9	−33.3	49.1	1	883	40.7	70.8	1	883
EPLS formula	All	−9.1	−48.8	30.6	27	351,449	35.9	63.0	36	190,580	−14.2	−56.1	27.7	27	351,449	33.8	56.6	36	190,580
EPLS formula	HIC	−13.1	−53.1	26.9	20	344,221	34.7	61.0	27	185,925	−14.5	−56.4	27.4	20	344,221	33.5	56.0	27	185,925
EPLS formula	LMIC	2.3	−36.4	40.9	7	7228	39.7	68.0	9	4655	−1.9	−39.1	35.2	7	7228	46.3	71.8	9	4655
Garwood formula	All	11.8	−38.9	62.5	2	1390	31.8	58.3	3	2461	10.7	−38.3	59.6	2	1390	30.6	57.3	3	2461
Garwood formula	HIC	9.2	−36.5	54.9	1	996	35.3	65.9	1	996	9.2	−36.5	54.9	1	996	35.3	65.9	1	996
Garwood formula	LMIC	14.4	−41.3	70.1	1	394	30.1	54.5	2	1465	14.4	−41.3	70.1	1	394	27.5	51.5	2	1465
Leffler formula	All	0.6	−42.8	44.0	10	270,349	32.6	57.4	11	271,373	−1.5	−40.1	37.2	10	270,349	43.8	69.3	11	271,373
Leffler formula	HIC	−3.7	−45.6	38.2	8	269,130	35.9	62.0	8	269,130	−1.5	−40.1	37.0	8	269,130	43.9	69.5	8	269,130
Leffler formula	LMIC	17.8	−31.6	67.1	2	1219	23.7	45.2	3	2243	11.8	−36.7	60.3	2	1219	27.8	52.3	3	2243
Luscombe formula	All	8.2	−34.5	50.9	15	237,568	34.0	58.1	21	275,841	0.4	−41.5	42.4	15	237,568	38.7	67.1	21	275,841
Luscombe formula	HIC	4.7	−36.3	45.8	11	231,881	38.3	64.7	15	269,554	0.2	−41.6	41.9	11	231,881	39.0	68.0	15	269,554
Luscombe formula	LMIC	17.7	−29.5	65.0	4	5687	23.1	46.2	6	6287	11.8	−33.7	57.2	4	5687	27.1	54.2	6	6287
Nelson formula	All	4.7	−33.6	42.9	9	134,315	36.8	61.9	10	135,339	−9.1	−38.6	20.5	9	134,315	42.0	73.7	10	135,339
Nelson formula	HIC	−6.7	−39.3	25.9	4	128,844	36.6	59.7	5	129,173	−9.8	−38.0	18.3	4	128,844	42.1	73.9	5	129,173

Table 3 Weight estimation meta-analysis summary data, showing both fixed effects (FE) and random effects (RE) data *(Continued)*

		Random effects outcomes									Fixed effects outcomes								
		MPE	LLOA	ULOA	Studies	*N*	PW10	PW20	Studies	*N*	MPE	LLOA	ULOA	Studies	*N*	PW10	PW20	Studies	*N*
Nelson formula	LMIC	13.8	−29.0	56.5	5	5471	37.0	63.7	5	6166	8.9	−30.3	48.0	5	5471	40.9	70.1	5	6166
Park formula	All	5.9	−35.4	47.2	2	125,170	36.9	64.3	4	126,788	1.0	−34.8	36.8	2	125,170	41.5	72.4	4	126,788
Park formula	HIC	0.9	−34.8	36.6	1	124,095	43.4	74.2	2	124,332	0.9	−34.8	36.6	1	124,095	41.7	72.8	2	124,332
Park formula	LMIC	10.9	−35.9	57.7	1	1075	30.5	54.5	2	2456	10.9	−35.9	57.7	1	1075	29.4	53.0	2	2456
Shann formula	All	−2.7	−49.1	43.8	7	107,890	35.8	60.5	9	111,266	−8.5	−41.9	24.9	7	107,890	40.2	70.9	9	111,266
Shann formula	HIC	−8.3	−53.2	36.5	4	106,150	37.2	62.1	5	108,455	−8.7	−41.6	24.3	4	106,150	40.3	71.1	5	108,455
Shann formula	LMIC	4.9	−43.7	53.5	3	1740	34.1	58.9	4	2811	3.4	−45.5	52.2	3	1740	35.5	61.2	4	2811
Theron formula	All	33.2	−47.9	114.2	5	4606	17.6	38.4	8	6362	29.4	−65.2	124.0	5	4606	18.8	13.5	8	6362
Theron formula	HIC	28.4	−54.5	111.3	3	3272	20.0	44.6	4	3577	27.8	−73.0	128.6	3	3272	18.8	13.5	4	3577
Theron formula	LMIC	40.3	−38.0	118.6	2	1334	15.2	32.2	4	2785	33.3	−43.7	110.4	2	1334	18.7	13.5	4	2785
Parental estimate	All	−3.6	−27.3	20.2	3	561	69.8	87.1	10	3070	−3.2	−24.9	18.4	3	561	78.1	89.8	10	3070
Parental estimate	HIC	−3.6	−27.3	20.2	3	561	68.0	87.1	9	2475	−3.2	−24.9	18.4	3	561	76.4	89.8	9	2475
Parental estimate	LMIC				0		85.2		1	595				0		85.2	595	1	
Healthcare provider guesses	All	−2.5	−51.7	46.8	3	547	35.6	59.3	5	1719	−2.8	−49.1	43.4	3	547	40.3	58.1	5	1719
Healthcare provider guesses	HIC	−2.5	−51.7	46.8	3	547	35.6	59.3	5	1719	−2.8	−49.1	43.4	3	547	40.3	58.1	5	1719
Healthcare provider guesses	LMIC				0				0					0				0	
Finger counting	All				0		53.8	83.4	2	4385	0.1	−34.0	34.2	0	4178	49.1	80.0	2	4385
Finger counting	HIC				0		53.8	83.4	2	4385	0.1	−34.0	34.2	0	4178	49.1	80.0	2	4385
Finger counting	LMIC				0				0		0.1	−34.0	34.2	0	4178			0	
Traub-Johnson formula	All	−0.8	−29.5	27.9	4	3561	45.9	74.8	4	3561	−3.2	−34.3	27.9	4	3561	48.8	75.2	4	3561
Traub-Johnson formula	HIC	−7.1	−35.9	21.7	2	2126	46.6	74.9	2	2126	−7.1	−37.4	23.2	2	2126	45.7	71.3	2	2126
Traub-Johnson formula	LMIC	5.5	−23.2	34.1	2	1435	45.2	74.7	2	1435	2.6	−26.4	31.5	2	1435	53.4	81.1	2	1435
Traub-Kichen formula	All	−2.0	−31.0	27.1	4	3455	45.4	74.4	4	3455	−4.8	−36.5	26.8	4	3455	48.5	74.4	4	3455
Traub-Kichen formula	HIC	−8.5	−38.2	21.2	2	2121	45.5	72.9	2	2121	−9.1	−39.8	21.6	2	2121	45.4	69.6	2	2121
Traub-Kichen formula	LMIC	4.6	−23.8	33.0	2	1334	45.3	75.9	2	1334	2.0	−26.4	30.4	2	1334	53.6	82.1	2	1334
MAC formula	All	6.7	−23.3	36.6	4	458,019	29.6	52.8	6	459,120	4.6	−16.0	25.3	4	458,019	28.0	58.0	6	459,120
MAC formula	HIC	7.4	−25.8	40.5	3	4029	29.3	51.5	4	4353	9.7	−29.4	48.9	3	4029	34.0	60.4	4	4353
MAC formula	LMIC	4.6	−15.8	25.0	1	453,990	30.1	58.0	2	454,767	4.6	−15.8	25.0	1	453,990	27.9	58.0	2	454,767
DWEM	All	−1.1	−26.3	24.1	2	884	57.8	87.6	3	1142	−0.9	−26.2	24.4	2	884	57.7	88.0	3	1142
DWEM	HIC	−1.1	−26.3	24.1	2	884	57.8	87.6	3	1142	−0.9	−26.2	24.4	2	884	57.7	88.0	3	1142
DWEM	LMIC				0				0					0				0	

Table 3 Weight estimation meta-analysis summary data, showing both fixed effects (FE) and random effects (RE) data *(Continued)*

		Random effects outcomes									Fixed effects outcomes								
		MPE	LLOA	ULOA	Studies	*N*	PW10	PW20	Studies	*N*	MPE	LLOA	ULOA	Studies	*N*	PW10	PW20	Studies	*N*
PAWPER tape	All	−0.2	−16.3	16.0	5	15,913	78.0	96.6	7	15,159	−1.0	−18.4	16.4	5	15,913	81.3	97.8	7	15,159
PAWPER tape	HIC	−0.8	−19.3	17.7	3	15,160	73.8	95.2	5	14,406	−1.0	−18.7	16.6	3	15,160	80.9	97.8	5	14,406
PAWPER tape	LMIC	0.8	−11.8	13.4	2	753	88.6	98.6	2	753	0.6	−12.0	13.3	2	753	88.7	98.7	2	753
Mercy method	All	−1.7	−19.5	16.2	8	5443	70.9	95.3	9	5642	−1.0	−19.7	17.7	8	5443	73.7	96.1	9	5642
Mercy method	HIC	−3.1	−20.9	14.8	5	4296	69.4	96.5	6	4495	−1.5	−20.2	17.1	5	4296	73.8	97.0	6	4495
Mercy method	LMIC	0.7	−17.2	18.7	3	1147	73.8	94.1	3	1147	0.9	−17.3	19.2	3	1147	73.2	93.6	3	1147
Wozniak method	All	−3.8	−36.1	28.5	2	628	72.1		2		−3.8	−36.3	28.6	2	628	74.9		2	1405
Wozniak method	HIC				0				0					0				0	
Wozniak method	LMIC	−3.8	−36.1	28.5	2	628	72.1		2		−3.8	−36.3	28.6	2	628	74.9		2	1405
Broselow tape	All	−1.5	−29.1	26.2	36	625,559	55.6	81.2	53	779,711	2.1	−20.6	24.9	36	625,559	61.7	91.2	53	779,711
Broselow tape	HIC	−3.9	−32.0	24.2	25	164,714	56.7	83.4	41	316,853	−4.7	−27.9	18.6	25	164,714	60.2	91.4	41	316,853
Broselow tape	LMIC	4.4	−25.9	34.7	11	462,079	52.0	76.3	12	462,530	4.6	−16.0	25.2	11	462,079	62.7	78.5	12	462,530

Data for the whole pooled sample as well as pooled data for high-income country (HIC) and low- and middle-income country (LMIC) populations are shown separately. There were very few substantial differences between the fixed effects and random effects analyses, which were not substantial enough to affect the overall outcomes. The number of studies in each pooled sample, as well as the number of data points is shown. A positive mean percentage error indicates an overestimation of weight, while a negative value indicates an underestimation of weight. Abbreviations: *MPE* mean percentage error, *LLOA* lower limit of agreement, *ULOA* upper limit of agreement, *PW10* percentage of weight estimates within 10% of actual weight, *PW20* percentage of weight estimates within 20% of actual weight

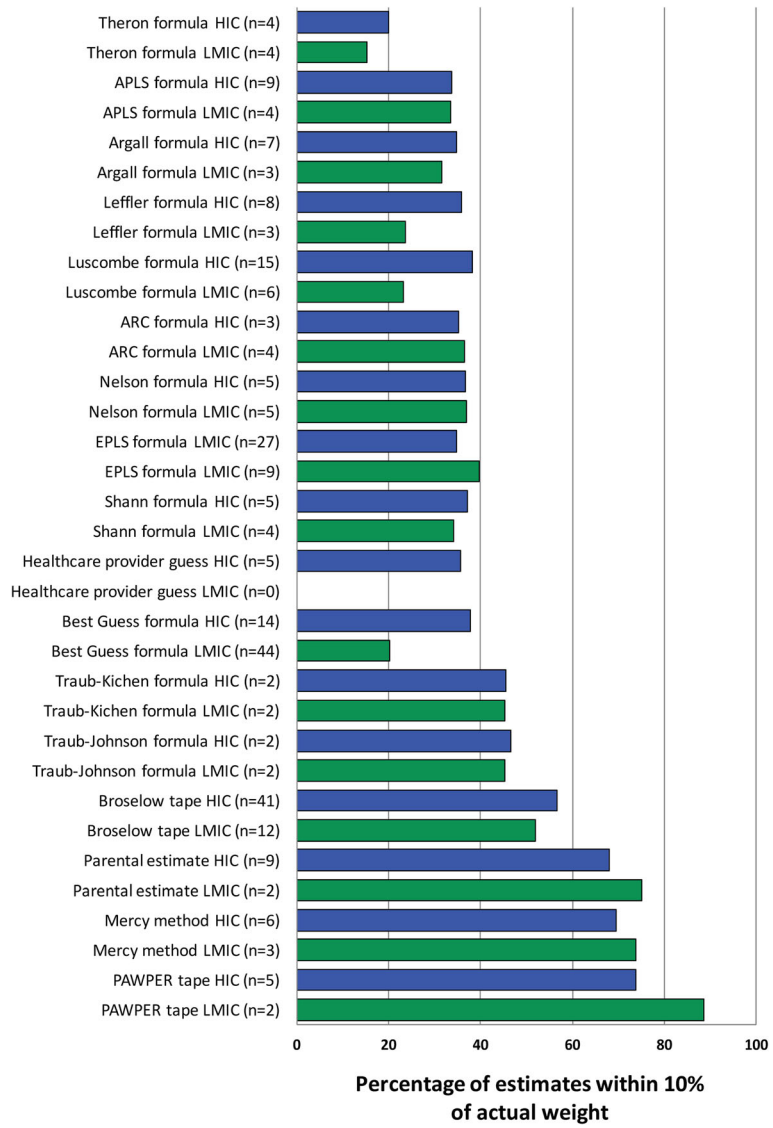


Fig. 3 A bar chart showing the accuracy data of the major weight estimation systems evaluated

personnel communicate a child's age during transport to hospital [35]. However, their use presupposes that a child's correct age is known, that the formula is remembered correctly and that the arithmetic is performed accurately. Memory is capricious in emergencies, however, and increased stress causes errors even in calculating simple formulas [44]. The benefits of the formulas are unlikely to mitigate for their very poor accuracy [11].

Many studies have shown age-based formulas to underestimate weight in first-world populations [45–47], but studies in low- and middle-income countries have shown a significant, potentially dangerous overestimation of weight by the same formulas [48–50]. In this meta-analysis, this was confirmed, with no age-based formula performing well in any population, but the overestimation of weight in low- and middle-income populations was significant and potentially unsafe. Even the use of habitus-modified age-formulas has failed to produce an improvement in accuracy to the degree of accuracy seen with length-based habitus-modified systems, as this modification still does not account for variations in length-for-age [11, 51].

This futility of age-based weight estimation can be perfectly summed up: "Accurate paediatric weight estimation by age: mission impossible" [27]. The unavoidable conclusion is that age-based formulas should no longer be used and clinicians that manage children should ensure that a better weight-estimation system is available for use during emergency care [11, 47, 52].

Table 4 Subgroup data for each weight estimation system

System	Subgroup		Mean percentage error	Lower limit of agreement	Upper limit of agreement	Number of children (number of studies)	PW10	Number of children (number of studies)
APLS formula (new)	Age < 1	FE	19.4	−30.7	69.4	5388 (2)	23.5	5700 (4)
		RE	8.2	−54.5	70.8		30.2	
	Age 1–5	FE	−6.0	−45.2	33.3	4438 (3)	37.3	8941 (5)
		RE	−8.2	−53.6	37.2		35.0	
	Age 6–12	FE	7.3	−50.7	65.2	2462 (3)	23.3	26,338 (5)
		RE	3.8	−56.2	63.8		29.9	
ARC formula	Age 1–5	FE	−10.1	−39.4	19.2	1415 (2)	43.4	1415 (2)
		RE	−9.8	−39.0	19.4		44.2	
	Age 6–10	FE	−12.4	−53.4	28.5	1405 (2)	33.2	1405 (2)
		RE	−12.5	−53.5	28.6		33.3	
Argall formula	Age 1–5	FE	3.8	−30.1	37.7	609 (1)	34.1	884 (2)
		RE					30.2	
	Age 6–10	FE	6.8	−36.9	50.5	639 (1)	27.8	741 (2)
		RE					25.0	
Best Guess formula	Age < 1	FE	2.4	−42.8	47.6	20,846 (5)	33.7	21,083 (7)
		RE	4.1	−43.9	52.1		38.4	
	Age 1–5	FE	5.0	−24.0	34.1	46,271 (7)	48.8	50,774 (9)
		RE	5.1	−20.0	30.2		44.1	
	Age 6–14	FE	6.7	−32.5	45.9	67,028 (6)	32.2	99,475 (9)
		RE	9.9	−37.8	57.5		31.4	
Chinese age-weight rule 1	Age 1–5	FE	−12.2	−51.3	26.8	1865 (2)	33.0	1865 (2)
		RE	−9.6	−45.8	26.6		35.5	
	Age 6–10	FE	−2.6	−50.1	44.9	1636 (2)	32.5	1636 (2)
		RE	−1.5	−47.5	44.5		32.1	
Chinese age-weight rule 2	Age 1–5	FE	−1.9	−33.9	30.0	1865 (2)	49.1	1865 (2)
		RE	−0.9	−32.7	30.9		48.0	
	Age 6–10	FE	−4.9	−52.1	42.3	1636 (2)	33.9	1636 (2)
		RE	−3.8	−49.5	41.9		33.3	
EPLS formula	Age 1–5	FE	−12.7	−43.4	18.0	91,652 (11)	37.9	96,077 (12)
		RE	−9.9	−42.6	22.9		41.5	
	Age 6–10	FE	−18.3	−61.7	25.0	101,742 (11)	29.0	116,988 (12)
		RE	−19.4	−63.8	24.9		28.1	
Leffler formula	Age < 1	FE	−5.7	−46.1	34.7	20,325 (2)	36.6	20,325 (2)
		RE	8.8	−42.8	60.4		27.7	
	Age 1–5	FE	4.2	−23.4	31.8	41,603 (3)	50.3	41,603 (3)
		RE	5.8	−25.4	37.0		43.8	
	Age 6–10	FE	−5.0	−39.1	29.1	64,426 (3)	42.0	64,426 (3)
		RE	−11.0	−45.6	23.6		31.9	
Luscombe formula	Age 1–5	FE	−3.3	−36.7	30.1	47,602 (4)	43.0	52,482 (7)
		RE	2.4	−32.5	37.3		37.7	
	Age 6–10	FE	1.2	−45.3	47.8	34,663 (4)	28.3	67,435 (7)
		RE	8.5	−35.7	52.6		28.9	
Nelson formula	Age < 1	FE	−8.1	−33.2	16.9	20,217 (2)	47.9	20,217 (2)
		RE	−3.5	−29.2	22.2		51.1	
	Age 1–5	FE	−9.1	−32.9	14.7	42,960 (3)	47.2	42,960 (3)
		RE	−5.7	−32.1	20.7		51.8	
	Age 6–10	FE	−10.6	−41.9	20.7	64,268 (2)	38.7	64,268 (2)
		RE	−9.4	−46.5	27.7		37.2	
Park formula	Age < 1		2.3	−42.8	47.4	19,854 (1)	33.5	19,854 (1)
	Age 1–5		−2.7	−28.0	22.6	40,612 (1)	55.2	40,612 (1)
	Age 6–10		2.7	−34.7	40.1	63,629 (1)	39.6	63,629 (1)

Table 4 Subgroup data for each weight estimation system *(Continued)*

	System	Subgroup		Mean percentage error	Lower limit of agreement	Upper limit of agreement	Number of children (number of studies)	PW10	Number of children (number of studies)
	Shann formula	Age 1–5	FE	−2.7	−28.1	22.7	41,221 (2)	55.1	41,221 (2)
			RE	−2.1	−31.0	26.8		53.0	
		Age 6–10	FE	−12.4	−46.2	21.4	64,268 (2)	34.8	64,268 (2)
			RE	−11.8	−50.5	26.9		35.0	
	Theron formula	Age 1–5		8.2	−27.1	43.5	609 (1)	36.8	609 (1)
		Age 6–10		30.7	−12.6	74.0	639 (1)	13.5	639 (1)
Dual length- and habitus based methods	Devised weight estimating method	<10 kg	FE	4.0	−21.9	29.9	221 (2)	52.7	221 (2)
			RE	3.7	−21.5	28.9		52.5	
		10-25 kg	FE	1.2	−17.7	20.1	232 (2)	70.2	232 (2)
			RE	1.3	−17.8	20.3		69.6	
		>25 kg	FE	−4.9	−31.9	22.2	332 (2)	54.0	431 (2)
			RE	−4.4	−30.5	21.7		53.5	
	PAWPER tape	<10 kg	FE	3.3	−13.8	20.4	765 (4)	75.6	783 (5)
			RE	2.7	−10.7	16.1		71.6	
		10-25 kg	FE	−0.2	−13.7	13.2	1022 (4)	86.2	1010 (4)
			RE	0.5	−11.2	12.2		85.5	
		>25 kg	FE	−2.4	−21.7	16.9	996 (4)	77.1	1070 (5)
			RE	−0.5	−17.0	16.0		76.8	
	Mercy Method	<10 kg	FE	−1.3	−25.1	22.6	103 (2)	62.9	121 (3)
			RE	−3.8	−25.3	17.7		60.0	
		10-25 kg	FE	−3.9	−20.4	12.6	218 (2)	74.3	296 (3)
			RE	−4.4	−20.0	11.2		73.0	
		>25 kg	FE	−4.9	−19.7	9.9	311 (2)	71.1	414 (3)
			RE	−4.3	−18.2	9.6		73.4	
	Wozniak method	<10 kg	FE	−21.7	−63.7	20.3	103 (2)	29.1	103 (2)
			RE	−23.7	−62.4	15.0		23.5	
		10-25 kg	FE	0.8	−31.9	33.5	218 (2)	69.5	218 (2)
			RE	0.8	−31.6	33.1		68.0	
		>25 kg	FE	−1.3	−19.8	17.2	311 (2)	78.8	311 (2)
			RE	−0.7	−17.6	16.2		81.3	
	Broselow tape	<10 kg	FE	−4.4	−29.3	20.5	26,327 (13)	55.2	27,345 (19)
			RE	−1.2	−19.7	17.4		53.0	
		10-25 kg	FE	−3.8	−22.8	15.1	51,915 (15)	66.4	57,102 (19)
			RE	−1.3	−18.9	16.3		60.7	
		>25 kg	FE	−5.3	−29.2	18.6	72,803 (14)	59.7	97,639 (22)
			RE	−2.8	−23.9	18.2		49.9	
Other	MAC formula	Age 1–5	FE	23.1	−8.3	54.4	1618 (2)	17.2	1618 (2)
			RE	21.3	−9.0	51.5		20.1	
		Age 6–10	FE	0.1	−39.8	40.0	1882 (2)	55.5	1882 (2)
			RE	0.0	−37.3	37.4		55.4	

The outcome data for the pooled data (not separated into high-income and low- and middle-income populations) is shown with both random effects (RE) and fixed effects (FE) results

Length-based weight estimation

Every length-based system performed better than every age-based system in this study. This supports the argument that length-based weight estimation is more biologically valid than age-based estimation [10]. No length-based system achieved the acceptable outcome benchmark, however.

The two length-based formulas were originally designed to predict ideal body weight in children, but they have been used, albeit incorrectly, to estimate total body weight. The addition of a habitus-modification to these formulas has been shown to increase their performance significantly, to the same level of accuracy as the other two-dimensional systems [11]. The use of these formulas in this way shows potential, especially if used with a mobile phone app, and requires further investigation.

Although there are at least seven length-only weight-estimation tapes, only the Broselow tape has been

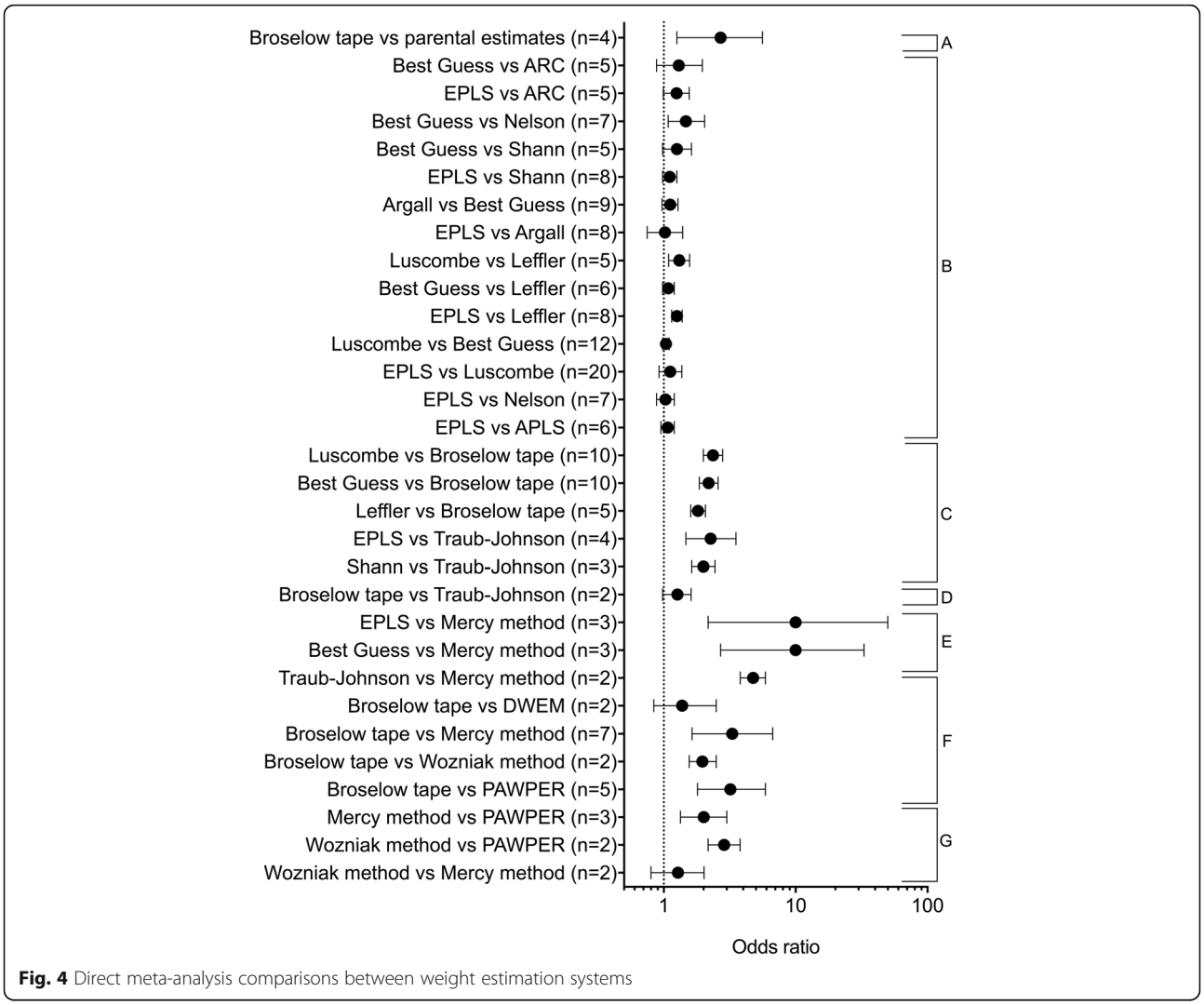


Fig. 4 Direct meta-analysis comparisons between weight estimation systems

extensively studied, while the Blantyre tape, the Sandell tape and the Handtevy tape have been evaluated only in single, small studies [53–55]. The Broselow tape, like other one-dimensional length-based systems, is vulnerable to error based on individual variations of weight-for-length (differences in body habitus) [56–59]. Some authors have questioned whether the tape is still valid given the increase in prevalence of overweight and obese children and may result in the "under-resuscitation of children" [33]. Although the manufacturer recommends modifying weight estimation up a colour zone in overweight children, to reduce this underestimation of weight, this has never been formally studied and still needs to be verified [60, 61]. However, while studies in high-income countries have demonstrated an overall underestimation of weight, studies in low- and middle-income countries have mostly shown an overestimation of weight, potentially to a dangerous degree in some populations (if drug doses were to be computed from those weights) [56, 57, 62]. Since length-based weight estimation is advocated by major, international advanced life support organisations and, since these systems are insufficiently accurate, this recommendation needs to be reconsidered and researched further [43, 63].

Two-dimensional (dual length- and habitus-based) weight estimation

The two-dimensional systems were far superior in accuracy to the one-dimensional age- and length-based systems. The accuracies of the Mercy method and the PAWPER tape in the meta-analysis were excellent, each with a PW10 of above 70% in both over- and undernourished populations. This finding was confirmed in individual studies, with no study reporting a one-dimensional system to be more accurate than a two-dimensional method. The direct meta-analysis comparisons showed that the PAWPER and Mercy

methods were significantly more accurate than the other systems, with the PAWPER tape outperforming the Mercy method in the two studies in which they were both evaluated.

All weight estimation systems have limitations, however. The Mercy method, like all other weight estimation systems was vulnerable to human factor errors in undertrained users [64]. It also has shown considerable variation in accuracy between individual assessors [19]. The functioning of the Mercy system in emergencies still needs to be evaluated– this is of concern as one of the poorest performances of the Mercy method was in a study which measured children in the supine position, as it might be used in an emergency [4, 20]. The PAWPER system was shown to be very accurate in two South African studies, one Australian study and one study based on NHANES data from the USA [5, 20–22, 31]. It was somewhat less accurate in two American studies with very obese populations, mostly because of difficulties in assessing body habitus, however [13, 65]. Although the tape's length-based measurements are objective and simple to perform, assessment of body habitus is more subjective and dependent on training and experience [66]. This will need to be researched further to explore more standardised and objective ways of assessing habitus.

The Devised Weight Estimating Method (DWEM), the Yamamoto obesity icon system, the Wozniak system and habitus-modified Traub-Johnson and Traub-Kichen formulas have all been shown to be significantly more accurate than length-based methods, but have not yet been sufficiently studied [11, 12, 14, 67].

Estimates of weight by parents

The utility of parental estimates of their child's weight is dependent on the parent being willing to offer a weight estimate and being accessible to healthcare personnel at the time of the child's need for emergency care [26]. The accuracy of prediction is determined by whether the accompanying parent is the regular caregiver of the child and whether or not the child has had a recent measurement of weight by the parent or in the parent's presence [9]. A previous systematic review has suggested that parental estimates are the most accurate method for obtaining a weight, when it cannot be measured [9]. In this meta-analysis, parental estimates were statistically superior to the Broselow tape on direct comparison, but there were no paired data from which direct comparisons could be made with the two-dimensional systems. Only one previous study has compared the Mercy method with parental estimates, in which parental estimates were found to be more accurate [68]. This will require further research to clarify, especially the accuracy of parental estimates in populations of different socioeconomic status and the frequency of availability of parental estimates. Since parents might not always be available, especially in the prehospital environment, it would be prudent to always have an alternative method of estimation available.

Differences in weight estimation accuracy between different populations

This study showed a clear disparity in how the one-dimensional weight estimation systems performed in different populations. These differences were primarily as a result of differences in *bias*, however, while the underlying lack of *precision* within each population was similar. Thus, the variability between populations was similar to the within-population variability shown in even the most homogeneous populations. The significance of this is that, although recalibration of a system for a specific population might reduce the bias, the underlying variability and imprecision would not allow an acceptable degree of overall accuracy to be achieved. This was well shown in the study by Asskaryar et al. which failed to recalibrate the Broselow tape in an Indian population by manipulating the bias only [57]. The two-dimensional systems, with their enhanced methodology which accounts for habitus, have proven to be the closer to a universally applicable system by achieving a more uniform accuracy, both within and between populations.

Limitations

The limitations of this study are similar to what is expected from any meta-analysis of this nature [24]. The lack of data comparing parental estimates and the newer two-dimensional systems limited the comparisons between these systems. The under-reporting on subgroups of weight status also limited the ability to analyse the performance of weight estimation systems in children with habitus that deviated from the average—this would provide insight into how the systems might function in populations with a high prevalence of underweight or obese children (or both).

Conclusions

No evidence exists of an acceptable benchmark for weight estimation systems. An accuracy of at least PW10 > 70% and PW20 > 95% could be considered as a reference standard, since the length-based, habitus-modified systems have proven that this target is achievable across a wide range of populations.

The only weight-estimation systems that were found to be of acceptable accuracy were the two-dimensional length- and habitus-based systems. The PAWPER tape and the Mercy Method achieved an accuracy that surpassed all other methods. Wide discrepancies in the accuracy of the Broselow tape in different age groups and

different populations raise questions about its use. It may dangerously overestimate weight in children from low- and middle-income countries or poor communities. Without exception, the age-based formulas evaluated proved to be highly inaccurate, with a possibility for patient harm, especially in low- and middle-income countries. There is sufficient evidence to conclude that the use of age-based formulas should be discouraged.

Recommendations

Dual length- and habitus-based (two-dimensional) systems should be used for weight estimation in children because of superior accuracy to other systems (high quality evidence).

The Broselow tape or parental estimates of weight should be used for weight estimation in preference to age-based formulas and healthcare provider guesses (medium quality evidence).

Age-based formulas and healthcare provider guesses should not be used for weight estimation in children because of potential patient harm (high quality evidence).

Parental estimates should be used to estimate weight in preference to length-based and age-based systems (high quality evidence). There was insufficient evidence to provide a recommendation between the two-dimensional systems and parental estimates of weight.

Additional file

Additional file 1: Table S1. Data from each study included in the meta-analysis, including subgroup data where available. (DOCX 145 kb)

Additional file 2: Figure S1. (JPEG 3645 kb)

Abbreviations

APLS: Advanced Paediatric Life Support formula; ARC: Australian Resuscitation Council formula; BG: Best Guess formula; BMI: Body mass index (kg/m^2); BT: Broselow tape; CAWR: Chinese age-weight rule formula; DWEM: Devised weight-estimating method; EPLS: European Paediatric Life support formula; FE: Fixed effects; LOA: 95% limits of agreement; MAC: Mid-arm circumference; MPE: Mean percentage error; NCHS: National Centre for Health Statistics; OR: Odds ratio; PW10: Percentage of estimates within 10% of actual weight; PW20: Percentage of estimates within 20% of actual weight; RE: Random effects; RMSPE: Root mean squared percentage error (%); TJ: Traub-Johnson formula; TK: Traub-Kichen formula

Funding

No external funding for this manuscript.

Financial disclosure

MW is the developer of the PAWPER tape, but derives no financial or commercial benefit from it. LG and AB have no financial relationships relevant to this article to disclose.

Authors' contributions

MW conceptualised and designed the study, carried out the initial data collection, analysed the data, drafted the initial manuscript and approved the final manuscript as submitted. LG carried out the initial data collection, drafted revisions of the manuscript and approved the final manuscript as submitted. AB assisted with preparation of the manuscript and approved the final manuscript as submitted. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Competing interests

MW is the developer of the PAWPER tape. LG and AB have no potential conflicts of interest to disclose.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 28 June 2017 Accepted: 7 September 2017
Published online: 21 September 2017

References

1. Luscombe M, Owens B. Weight estimation in resuscitation: is the current formula still valid? Arch Dis Child. 2007;92(5):412–5.
2. Meguerdichian MJ, Clapper TC. The Broselow tape as an effective medication dosing instrument: a review of the literature. J Pediatr Nurs. 2012;27(4):416–20.
3. Luten R, Zaritsky A. The sophistication of simplicity... optimizing emergency dosing. Acad Emerg Med. 2008;15(5):461–5.
4. Abdel-Rahman SM, Ridge A, Kearns GL. Estimation of body weight in children in the absence of scales: a necessary measurement to insure accurate drug dosing. Arch Dis Child. 2014;99(6):570–4.
5. Wells M, Coovadia A, Kramer E, Goldstein L. The PAWPER tape: a new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. Resuscitation. 2013;84(2):227–32.
6. Hoyle JD, Davis AT, Putman KK, Trytko JA, Fales WD. Medication dosing errors in pediatric patients treated by emergency medical services. Prehospital Emergency Care. 2012;16(1):59–66.
7. Lammers R, Willoughby-Byrwa M, Fales W. Medication errors in prehospital management of simulated pediatric anaphylaxis. Prehospital Emergency Care. 2014;18(2):295–304.
8. Stevens AD, Hernandez C, Jones S, Moreira ME, Blumen JR, Hopkins E, et al. Color-coded prefilled medication syringes decrease time to delivery and dosing errors in simulated prehospital pediatric resuscitations: a randomized crossover trial. Resuscitation. 2015;96:85–91.
9. Young KD, Korotzer NC. Weight estimation methods in children: a systematic review. Ann Emerg Med. 2016;68(4):441–51. e10
10. Goldman J, Ojoo A, Abdel-Rahman S. Challenges in pediatric oral dosing. In: MacLeod S, Hill S, Koren G, Rane A, editors. Optimizing treatment for children in the developing world. Cham: Springer International Publishing; 2015. p. 33–44.
11. Wells M, Goldstein L, Bentley A. It is time to abandon age-based emergency weight estimation in children! A failed validation of 20 different age-based formulas. Resuscitation. 2017;116(7):73–83.
12. Yamamoto LG, Inaba AS, Young LL, Anderson KM. Improving length-based weight estimates by adding a body habitus (obesity) icon. Am J Emerg Med. 2009;27(7):810–5.
13. Chavez H, Peterson R, Lo K, Arel M. Weight estimation in an inner city Pediatric Emergency Department: the effect of obesity. Am J Emerg Med. 2015;33(10):1364–7.
14. Whitfield KC, Wozniak R, Pradinuk M, Karakochuk CD, Anabwani G, Daly Z, et al. Anthropometric measures are simple and accurate paediatric weight-prediction proxies in resource-poor settings with a high HIV prevalence. Arch Dis Child. 2017;102(1):10–6.
15. Abdel-Rahman SM, Ridge AL. An improved pediatric weight estimation strategy. Open Medical Devices Journal. 2012;4:87–97.
16. Jung JY, Kwak YH, Kim DK, Suh D, Chang I, Yoon C, et al. An assessment of the accuracy of a novel weight estimation device for children. Emerg Med J. 2017;34(3):163–169.
17. Tanner D, Negaard A, Huang R, Evans N, Hennes H. A prospective evaluation of the accuracy of weight estimation using the Broselow Tape in overweight and obese pediatric patients in the Emergency Department. Pediatr Emerg Care. 2016. doi:10.1097/PEC.0000000000000894. [Epub ahead of print].
18. Abdel-Rahman SM, Paul IM, James LP, Lewandowski A. Evaluation of the Mercy TAPE: performance against the standard for pediatric weight estimation. Ann Emerg Med. 2013;62(4):332–339.e6.

19. Abdel-Rahman SM, Ahlers N, Holmes A, Wright K, Harris A, Weigel J, et al. Validation of an improved pediatric weight estimation strategy. J Pediatr Pharmacol Ther. 2013;18(2):112–21.
20. Wells M, Goldstein L, Bentley A. A validation study of the PAWPER XL tape: accurate estimation of both total and ideal body weight in children up to 16 years of age. Trauma and Emergency Care. 2017;2(4):1–8.
21. Georgoulas V, Wells M. The PAWPER tape and the Mercy Method outperform other methods of weight estimation in children in South Africa. S Afr Med J. 2016;106(9):933–9.
22. O'Leary F, John-Denny B, McGarvey K, Hann A, Pegiazoglou I, Peat J. Estimating the weight of ethnically diverse children attending an Australian emergency department: a prospective, blinded, comparison of age-based and length-based tools including Mercy, PAWPER and Broselow. Arch Dis Child. 2017;102(1):46–52.
23. Wells M, Goldstein L, Bentley A. A systematic review and meta-analysis of the accuracy of weight estimation systems used in paediatric emergency care in developing countries. African Journal of Emergency Medicine. 2017; In Press
24. Williamson PR, Lancaster GA, Craig JV, Smyth RL. Meta-analysis of method comparison studies. Stat Med. 2002;21(14):2013–25.
25. Berman NG, Parker RA. Meta-analysis: neither quick nor easy. BMC Med Res Methodol. 2002;2:10.
26. Kelly A. Estimation of children's weight in medical emergencies. In: Preedy V, editor. Handbook of anthropometry: physical measures of human form in health and disease. New York: Springer-Verlag; 2012. p. 1151–61.
27. Marlow R, Lo D, Walton L. Accurate paediatric weight estimation by age: mission impossible? Arch Dis Child. 2011;96(Suppl 1):A1–2.
28. Park J, Kwak YH, Kim do K, Jung JY, Lee JH, Jang HY, et al. A new age-based formula for estimating weight of Korean children. Resuscitation. 2012;83(9):1129–34.
29. Ralston ME, Myatt MA. Weight estimation tool for children aged 6 to 59 months in limited-resource settings. PLoS One. 2016;11(8):e0159260.
30. Stewart D. Accuracy of the Broselow tape for estimating paediatric weight in two Australian Emergency Departments: University of Sydney; 2009.
31. Wells M, Goldstein L, Bentley A. Development and validation of a method to estimate bodyweight in critically ill children using length and mid-arm circumference measurements—the PAWPER XL-MAC system. S Afr Med J. 2017; In Press
32. Hashikawa A, Juhn Y, Homme J, Gardner B, Moore B. Does length-based resuscitation tape accurately place pediatric patients into appropriate color-coded zones? Pediatr Emerg Care. 2007;23(12):856–61.
33. Milne KW, Yasin A, Knight J, Noel D, Lubell R, Filler G. Ontario children have outgrown the broselow tape. Canadian Journal of Emergency Medicine. 2012;14(1):25–30.
34. Thompson MT, Reading MJ, Acworth JP. Best Guess method for age-based weight estimation in paediatric emergencies validation and comparison with current methods. Emergency Medicine Australasia. 2007;19:535–42.
35. Williams B, Boyle M, O'Meara P. Can undergraduate paramedic and nursing students accurately estimate patient age and weight? Prehospital and Disaster Medicine. 2010;25(2):171–7.
36. Greig A, Ryan J, Glucksman E. How good are doctors at estimating children's weight? J Accid Emerg Med. 1997;14:101–3.
37. Nieman CT, Manacci CF, Super DM, Mancuso C, Fallon WF Jr. Use of the Broselow tape may result in the underresuscitation of children. Acad Emerg Med. 2006;13(10):1011–9.
38. Rosenberg M, Greenberger S, Rawal A, Latimer-Pierson J, Thundiyil J. Comparison of Broselow tape measurements versus physician estimations of pediatric weights. Am J Emerg Med. 2011;29(5):482–8.
39. Lim CAE, Kaufman BJ, O'Connor J Jr, Cunningham SJ. Accuracy of weight estimates in pediatric patients by prehospital Emergency Medical Services personnel. Am J Emerg Med. 2013;31(7):1108–12.
40. Nash DB. Meet the enemy: he is us. Am J Med Qual. 2011;26(1 Suppl):4S.
41. Luten R. Error and time delay in pediatric trauma resuscitation: Addressing the problem with color-coded resuscitation aids. Surg Clin N Am. 2002; 82(2):303–14.
42. Winship C, Williams B, Boyle M. Which paediatric weight formula is best suited for the out-of-hospital field? Journal of Paramedic Practice. 2012;4(9):533–41.
43. Advanced Life Support Group. Advanced Paediatric Life Support: The Practical Approach (APLS). 5th ed. New York: BMJ Books - Publisher: John Wiley & Sons (Wiley-Blackwell); 2011.
44. Young TP, Chen BG, Kim TY, Thorp AW, Brown L. Finger counting: an alternative method for estimating pediatric weights. Am J Emerg Med. 2014;32(3):243–247.
45. Cattermole GN, Leung MP, So HK, Mak PS, Graham CA, Rainer TH. Age-based formulae to estimate children's weight in the emergency department. Emerg Med J. 2011;28(5):390–6.
46. Flannigan C, Bourke TW, Sproule A, Stevenson M, Terris M. Are APLS formulae for estimating weight appropriate for use in children admitted to PICU? Resuscitation. 2014;85(7):927–31.
47. Britnell S, Koziol-McLain J. Weight estimation in paediatric resuscitation: a hefty issue in New Zealand. Emergency Medicine Australasia. 2015;27(3):251–6.
48. Pollock L, Molyneux E, Jefferis O, Bunn J. Estimating children's weights: a developing country perspective. Arch Dis Child. 2007;92(10):938–9.
49. Batmanabane G, Kumar Jena P, Dikshit R, Abdel-Rahman SM. Using the Mercy method for weight estimation in Indian children. Global Pediatric Health. 2015; January-December:1-7.
50. Dicko A, Alhousseini ML, Sidibé B, Traoré M, Abdel-Rahman SM. Evaluation of the Mercy weight estimation method in Ouelessebougou, Mali. BMC Public Health. 2014;14:270.
51. Britnell S, Taylor S, Koziol-McLain J. Emergency weight estimation lookup tables for New Zealand children aged 5-10 years. Emergency Medicine Australasia. 2016;28(5):558–63.
52. Black K, Barnett P, Wolfe R, Young S. Are methods used to estimate weight in children accurate? Emerg Med. 2002;14(2):160–5.
53. Molyneux E, Brogan R, Mitchell G, Gove S. Children's weights: guess or measure by tape. Lancet. 1999;354:1616.
54. Allison N, Norton I. Emergency weight estimation in Aboriginal and Torres Strait Islander children in the Northern Territory: are the current methods accurate? Emergency Medicine Australasia. 2014;26(5):487–93.
55. Lowe CG, Campwala RT, Ziv N, Wang VJ. The Broselow and Handtevy Resuscitation Tapes: a comparison of the performance of pediatric weight prediction. Prehospital and Disaster Medicine. 2016;31(4):364–75.
56. Ramarajan N, Krishnamoorthi R, Strehlow M, Quinn J, Mahadevan SV. Internationalizing the Broselow tape: how reliable is weight estimation in Indian children. Acad Emerg Med. 2008;15(5):431–6.
57. Asskaryar F, Shankar R. An Indian pediatric emergency weight estimation tool: prospective adjustment of the Broselow tape. Int J Emerg Med. 2015;8(1):78.
58. Wells M, Goldstein L, Bentley A. Accuracy of weight estimation by the Broselow tape is substantially improved by including a visual assessment of body habitus. Pediatric Research. 2017. In Press. doi:10.1038/pr.2017.222.
59. Clark M, Ogunnyi A, Lewis R, Patel D. Is the Broselow tape valid in South Sudan, "the hungriest place on earth"? Ann Emerg Med. 2013;62(4):S98.
60. Broselow® Pediatric Emergency Tape 2011 [Acessed 25 Sept 2015, 2015]. Available from: https://www.armstrongmedical.com/index.cfm/go/product.detail/sec/3/ssec/14/fam/2371.
61. Deboer S, Seaver M, Broselow J. Color coding to reduce errors. Am J Nurs. 2005;105(8):68–71.
62. Clark MC, Lewis RJ, Fleischman RJ, Ogunniyi AA, Patel DS, Donaldson RI. Accuracy of the Broselow Tape in South Sudan, "The Hungriest Place on Earth". Acad Emerg Med. 2016;23(1):21–8.
63. Kleinman ME, Chameides L, Schexnayder SM, Samson RA, Hazinski MF, Atkins DL, et al. Pediatric advanced life support: 2010 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Pediatrics. 2010;126(5):e1361–99.
64. Abdel-Rahman A, Jacobsen R, Watts J, Doyle S, O'Malley D, Hefner T, et al. Comparative performance of pediatric weight estimation techniques: a human factor errors analysis. Pediatr Emerg Care. 2015;00:00.
65. Garcia CM, Meltzer JA, Chan KN, Cunningham SJ. A validation study of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape-a new weight estimation tool. J Pediatr. 2015;167(1):173–177.e1.
66. Goldstein L, Wells M. Fat is the new normal. J Pediatr. 2016;168:257.
67. Garland JS, Kishaba RG, Nelson DB, Losek JD, Sobocinski KA. A rapid and accurate method of estimating body weight. Am J Emerg Med. 1986;4(5):390–3.
68. Samerchua A, Suraseranivongse S, Komoltri C. A Comparison of pediatric weight estimation methods for emergency resuscitation. Pediatr Emerg Care. 2017. doi:10.1097/PEC.0000000000001137. [Epub ahead of print].
69. Traub SL, Kichen L. Estimating ideal body mass in children. Am J Hosp Pharm. 1983;40(1):107–10.

70. Lubitz DS, Seidel JS, Chameides L, Luten RC, Zaritsky AL, Campbell FW. A rapid method for estimating weight and resuscitation drug dosages from length in the pediatric age group. Ann Emerg Med. 1988;17(6):576–81.
71. Oakley P. Inaccuracy and delay in decision making in paediatric resuscitation, and a proposed reference chart to reduce error. Br Med J. 1988;297:817–9.
72. Losek JD, Garland J, Nelson DB. Body habitus + height = accurate weight estimate. Ann Emerg Med. 1989;18(6):709.
73. Haftel AJ, Khan N, Lev R, Schonfeld N. Hanging leg weight—a rapid technique for estimating total body weight in pediatric resuscitation. Ann Emerg Med. 1990;19(5):523–6.
74. Hughes G, Spoudeas H, Kovar IZ, Millington HT. Tape measure to aid prescription in paediatric resuscitation. Archives of Emergency Medicine. 1990;7(1):21–7.
75. Leffler S, Hayes M. Analysis of parental estimates of children's weights in the ED. Ann Emerg Med. 1997;30(2):167–70.
76. Dearlove CB, Dearlove O. Visual estimation of children's weights. Anaesthesia. 1999;54(12):1228–9.
77. Goldman RD, Buskin S, Augarten A. Parental estimates of their child's weight: accurate for resuscitation drug doses. Pediatr Emerg Care. 1999;15(1):19–21.
78. Harris M, Patterson J, Morse J. Doctors, nurses and parents are equally poor at estimating pediatric weights. Pediatr Emerg Care. 1999;15:17–8.
79. Kun W, Cheng K, Yuen M, Tung W. How good is the Broselow tape measurement for estimation of body weights in paediatric patients for application in Hong Kong? Hong Kong Journal of Paediatrics. 2000;5(1):25–30.
80. Carroll W, Jay N, Alexander J. Towards better weight estimation in the seriously ill child - a comparison of methods. Arch Dis Child. 2001; 84(Supp I):A12.
81. Vilke GM, Marino A, Fisher R, Chan TC. Estimation of pediatric patient weight by EMT-PS. J Emerg Med. 2001;21(2):125–8.
82. Hofer CK, Ganter M, Tucci M, Klaghofer R, Zollinger A. How reliable is length-based determination of body weight and tracheal tube size in the paediatric age group? The Broselow tape reconsidered. Br J Anaesth. 2002; 88(2):283–5.
83. Uesugi T, Okada N, Sakai K, Nishina K, Mikawa K, Shiga M. Accuracy of visual estimation of body height and weight in supine paediatric patients. Paediatr Anaesth. 2002;12(6):489–94.
84. Argall JA, Wright N, Mackway-Jones K, Jackson R. A comparison of two commonly used methods of weight estimation. Arch Dis Child. 2003; 88(9):789–90.
85. Potier K. Broselow tape or APLS formula to estimate weight in children. 2006. Available from: http://bestbets.org/bets/bet.php?id=64.
86. Hohenhaus SM, Frush KS. Pediatric patient safety: common problems in the use of resuscitative aids for simplifying pediatric emergency care. J Emerg Nurs. 2004;30(1):49–51.
87. Moore B, Hankins D, Shimmin S. Use of the length-based resuscitation tape for estimating pediatric weight in the pre-hospital setting. Prehospital Emergency Care. 2004;8(1):87.
88. Luscombe M. "Kid's aren't what they used to be": a study of paediatric patients' weights and their relationship to current weight estimation formulae. Br J Anaesth. 2005;95(4):578.
89. Theron L, Adams A, Jansen K, Robinson E. Emergency weight estimation in Pacific Island and Maori children who are large-for-age. Emergency Medicine Australasia. 2005;17:238–43.
90. Bavdekar SB, Sathe S, Jani P. Prediction of weight of Indian children aged up to two years based on foot-length: implications for emergency areas. Indian Pediatr. 2006;43(2):125–30.
91. Varghese A, Vasudevan VK, Lewin S, Indumathi CK, Dinakar C, Subba Rao SD. Do the length-based (Broselow®) tape, APLS, Argall and Nelson's formulae accurately estimate weight of Indian children? Indian Pediatr. 2006;43(10):889–94.
92. Anderson M. Underestimation potentially better than overestimation. Arch Dis Child. 2007;92(10):937.
93. Dieckmann RA. The dilemma of paediatric drug dosing and equipment sizing in the era of patient safety. Emergency Medicine Australasia. 2007;19:490–3.
94. DuBois D, Baldwin S, King WD. Accuracy of weight estimation methods for children. Pediatr Emerg Care. 2007;23(4):227–30.
95. Im T, Kang B, Choi H, Won T, Yeon J. A study on the accuracy of weight estimation using the Broselow tape in Korean children. Ann Emerg Med. 2007;50(3):S40.
96. Jang HY, Shin SD, Kwak YH. Can the Broselow tape be used to estimate weight and endotracheal tube size in Korean children? Acad Emerg Med. 2007;14(5):489–91.
97. Kelly AM, Kerr D, Clooney M, Krieser D, Nguyen K. External validation of the best Guess formulae for paediatric weight estimation. Emergency Medicine Australasia. 2007;19(6):543–6.
98. Krieser D, Nguyen K, Kerr D, Jolley D, Clooney M, Kelly AM. Parental weight estimation of their child's weight is more accurate than other weight estimation methods for determining children's weight in an emergency department? Emerg Med J. 2007;24(11):756–9.
99. Luten R, Kahn N, Wears R, Kissoon N. Predicting endotracheal tube size by length in newborns. J Emerg Med. 2007;32(4):343–7.
100. Nguyen K, Krieser D, Kerr D, Jolley D, Clooney M, Kelly AM. Failed validation of the Argall weight formula for estimating children's weight in an Australian Emergency Department. Acad Emerg Med. 2007;14(5):486–8.
101. Patel A, Zikos V, Rzechula K, McQuillen K. The accuracy of pediatric weight in the Emergency Department: do caregiver weight estimates measure up? Ann Emerg Med. 2007;50(3, Supplement):S123–4.
102. Tinning K, Acworth J. Make your Best Guess: an updated method for paediatric weight estimation in emergencies. Emergency Medicine Australasia. 2007;19:528–34.
103. Zikos V, Patel A, Rzechula K, K M. The accuracy of pediatric weight in the emergency department: do caregiver weight estimates measure up? Ann Emerg Med. 2007;50(3):S123–4.
104. Gardner S, Haber R. Children's weight "guesstimates": could we? Ann Emerg Med. 2008;51(4):484–5.
105. Wells M, Kramer E. Optimizing emergency drug dosing in children. Acad Emerg Med. 2008;15(12):1325.
106. Zink K, Bohn M, Hood T, Topp RV, Berger J. A comparison of methods of assessing patient body weight in the pediatric emergency department. Kentucky Nurse. 2008;56(3):12.
107. Anstett D, Bawden J, Moylette E. Does the Broselow tape accurately estimate the weight of healthy Irish children? Canadian Journal of Emergency Medicine. 2009;11:289.
108. Cattamanchi S, Banala S, RV T. Assessing competency of the Broselow-Luten pediatric resuscitation tape: a prospective, cross sectional, analytical study of 15,000 Indian school children. Ann Emerg Med. 2009;54(3):S130–1.
109. Cattermole GN, Leung M, Mak P, So HK, Graham CA, Rainer TH. Children's weights correlate more strongly with mid-arm circumference (MAC) than with age, height or foot-length. J Emerg Med. 2009;37(2):228–9.
110. Partridge RL, Abramo TJ, Haggarty KA, Hearn R, Sutton KL, An AQ, et al. Analysis of parental and nurse weight estimates of children in the pediatric emergency department. Pediatr Emerg Care. 2009;25(12):816–8.
111. Paw R, Majeed MA. Emergency paediatric weight estimation. Does the APLS formula hold true? J Emerg Med. 2009;37(2):231.
112. Sandell JM, Charman SC. Can age-based estimates of weight be safely used when resuscitating children? Emerg Med J. 2009;26(1):43–7.
113. So TY, Farrington E, Absher RK. Evaluation of the accuracy of different methods used to estimate weights in the pediatric population. Pediatrics. 2009;123(6):e1045–51.
114. Bicer S, Oguzhan A, Vitrinel A. Evaluation of the weight estimation formulae in childhood used frequently. Türkiye Klinikleri Pediatri Dergisi. 2010;19(3): 197–202.
115. Casey J, Borland M. Best Guess method: a further external validation study and comparison with other methods. Emergency Medicine Australasia. 2010;22(1):68–74.
116. Cattermole GN, Leung PY, Mak PS, Graham CA, Rainer TH. Mid-arm circumference can be used to estimate children's weights. Resuscitation. 2010;81(9):1105–10.
117. Lulic I, Kovic I. Comparing methods for weight estimation of children: a preliminary study. Resuscitation. 2010;81(2, Supplement):S77.
118. Rosenberg M, Thundiyil J, Greenberger S, Rawal A, Latimer-Pierson J. Does physician estimates of pediatric patient weights lead to inaccurate medication dosages. Ann Emerg Med. 2010;56(3 (Suppl September 2010):S47.
119. Bourdeau S, Copeland J, Milne WK. Accuracy of the Broselow tape in estimating the weight of First Nations children. Canadian Journal of Rural Medicine. 2011;16(4):121–5.
120. Cattermole GN, Graham CA, Rainer TH. Age-based weight prediction is imprecise. Emerg Med J. 2011;29(1):82–3.
121. Costelloe C, Montgomery AA, Redmond NM, Fletcher M, Hollinghurst S, Peters TJ, et al. Medicine dosing by weight in the home: can parents accurately weigh preschool children? A method comparison study. Arch Dis Child. 2011;96(12):1187–90.

122. Geduld H, Hodkinson PW, Wallis LA. Validation of weight estimation by age and length based methods in the Western Cape. South Africa population Emergency Medicine Journal. 2011;28(10):856–60.
123. Huybrechts I, Himes JH, Ottevaere C, De Vriendt T, De Keyzer W, Cox B, et al. Validity of parent-reported weight and height of preschool children measured at home or estimated without home measurement: a validation study. BMC Pediatr. 2011;11:63.
124. Kelly AM, Nguyen K, Krieser D. Validation of the Luscombe weight formula for estimating children's weight. Emergency Medicine Australasia. 2011; 23(1):59–62.
125. Knight JC, Nazim M, Riggs D, Channel J, Mullet C, Vaughan R, et al. Is the Broselow tape a reliable indicator for use in all pediatric trauma patients?: a look at a rural trauma center. Pediatr Emerg Care. 2011;27(6):479–82.
126. Luscombe MD, Owens BD, Burke D. Weight estimation in paediatrics: a comparison of the APLS formula and the formula 'Weight=3(age)+7'. Emerg Med J. 2011;28(7):590–3.
127. Wells M. The PAWPER tape: a more accurate form of tape-based weight estimation. Sanguine. 2011;1(2):4–6.
128. Ali K, Sammy I, Nunes P. Is the APLS formula used to calculate weight-for-age applicable to a Trinidadian population? BMC Emergency Medicine. 2012;12:9.
129. Cantle F, Cattermole G, Graham C, Rainer T. Validation of the Chinese age weight rules and the mid-arm circumference method to estimate children's weight. Dublin: International Conference on Emergency Medicine; 2012.
130. Elgie LD, Williams AR. Using age on clothes size label to estimate weight in emergency paediatric patients. Eur J Emerg Med. 2012;19(5):338–40.
131. Garwood J, McEwan A. Estimating paediatric weight: a new formula. Paediatric Anaesthesia and Intensive Care. 2012;29(Suppl 50):153–4.
132. Heyming T, Bosson N, Kurobe A, Kaji AH, Gausche-Hill M. Accuracy of paramedic Broselow tape use in the prehospital setting. Prehospital Emergency Care. 2012;16(3):374–80.
133. Seddon C, Lockitt L, Dhanjal S, Eisenhut M. Validation of advanced paediatric life support formulas for weight calculation in a multiethnic population. International Scholarly Research Notices: Pediatrics. 2012; 2012:869634.
134. Sinha M, Lezine MW, Frechette A, Foster KN. Weighing the pediatric patient during trauma resuscitation and its concordance with estimated weight using Broselow Luten emergency tape. Pediatr Emerg Care. 2012;28(6):544–7.
135. Trakulsrichai S, Boonsri C, Chatchaipun P, Chunharas A. Accuracy of three methods used for Thai children's body weight estimation. J Med Assoc Thail. 2012;95(9):1194–9.
136. Wozniak R. The evaluation of potential weight-estimation methods in a primarily HIV positive cohort in Botswana for use in resource limited settings: The University of British Columbia; 2012.
137. Akabarian S, Vahedparast H, Ravanipour M, Mirzaei K. Comparison of Broselow tape measurements versus mother estimations of pediatric weights. Iranian South Medical Journal. 2013;16(2):138–43.
138. Cattermole GN, Leung PY, Graham CA, Rainer TH. Too tall for the tape: the weight of schoolchildren who do not fit the Broselow tape. Emerg Med J. 2013; [epub ahead of print]
139. Graves L, Chayen G, Peat J, O'Leary F. A comparison of actual to estimated weights in Australian children attending a tertiary children's hospital, using the original and updated APLS, Luscombe and Owens, Best Guess formulae and the Broselow tape. Resuscitation. 2014;85(3):392–6.
140. Hegazy M, Taher E. Validating a new formula for weight estimation in pediatric cancer patients. International Research Journal of Medicine and Medical Sciences. 2013;1(1):34–9.
141. House DR, Ngetich E, Vreeman RC, Rusyniak DE. Estimating the weight of children in Kenya: do the Broselow tape and age-based formulas measure up? Ann Emerg Med. 2013;61(1):1–8.
142. Loo PY, Chong SL, Lek N, Bautista D, Ng KC. Evaluation of three paediatric weight estimation methods in Singapore. J Paediatr Child Health. 2013; 49(4):E311–6.
143. Suh D, Kwak Y, Kim D, Jeong J. Estimating weights using the Broselow tape in Korea: comparison of old and new versions of the Broselow tape use. Ann Emerg Med. 2013;62(4):S109.
144. Ackwerh R, Lehrian L, Nafiu OO. Assessing the accuracy of common pediatric age-based weight estimation formulae. Anesth Analg. 2014; 118(5):1027–33.
145. Chiengkriwate P, Donnapee R, Geater A. The accuracy of the Broselow tape in the weight estimation of Thai children. Asian Biomedicine. 2014;8(6):799–807.
146. Eke C, Ubesie A, Ibe B. Comparison of actual (measured) weights and heights with the standard formula methods of estimation among children in Enugu. Nigerian Journal of Pediatrics. 2014;41:307–11.
147. Erker CG, Santamaria M, Moellmann M. Size does matter—age-related weight estimation in "tall n' thin" and "tiny n' thick" children and a new habitus-adapted alternative to the EPLS-formula. Resuscitation. 2014;85(9):1174–8.
148. Omisanjo AO, Orimadegun AE, Akinbami FO. Accuracy of Nelson and best Guess formulae in estimation of weights in Nigerian children population. Annals of Ibadan Postgraduate Medicine. 2014;12(2):80–8.
149. Badeli H, Hashemian H, Nazari N, Rad AH. The percentage of error of different weight estimation methods toward actual weight in children admitted to 17 Shahrivar Hospital. International Journal of Preventive Medicine. 2015;6:13.
150. Khouli M, Ortiz MI, Romo-Hernández G, Martínez-Licona D, Stelzner SM. Use of the Broselow tape in a Mexican Emergency Department. J Emerg Med. 2015;48(6):660–6.
151. Skrobo D, Kelleher G. CORKSCREW 2013 CORK study of children's realistic estimation of weight. Emerg Med J. 2015;32(1):32–5.
152. Talib NJ, Rahm G, Abdel-Rahman SM. Estimating weight in children with Down syndrome. Global Pediatric Health. 2015;2:2333794X14568450.
153. Young TP, Washington O, Flanery A, Guptill M, Reibling ET, Brown L, et al. Comparison of the finger counting method, the Broselow tape and common weight estimation formulae in Filipino children after Typhoon Haiyan. Emergency Medicine Australasia. 2015;27(3):239–44.
154. AlHarbi TM, AlGarni A, AlGamdi F, Jawish M, Wani TA, Abu-Shaheen AK. The accuracy of Broselow Tape weight estimate among pediatric population. Biomed Res Int. 2016;2016:7042947.
155. Aliyu I, Abdulsalam M. Is the Broselow tape and APLS formula applicable to Nigerian children? The Kano experience. Tropical Journal of Medical Research. 2016;19(1):20–3.
156. Carasco CF, Fletcher P, Maconochie I. Review of commonly used age-based weight estimates for paediatric drug dosing in relation to the pharmacokinetic properties of resuscitation drugs. Br J Clin Pharmacol. 2016;81(5):849–56.
157. Chassee T, Reischmann D, Mancera M, Hoyle JD. Emergency medical dispatchers can obtain accurate pediatric weights from 9-1-1 callers. Prehospital Emergency Care. 2016;20(6):808–14.
158. Jung J, Kwon H, Choi Y. Methods of weight estimation in pediatric resuscitation. Pediatric Emergency Medicine Journal. 2016;3(1):9–14.
159. Mishra DG, Kole T, Nagpal R, Smith JP. A correlation analysis of Broselow Pediatric Emergency Tape-determined pediatric weight with actual pediatric weight in India. World J Emerg Med. 2016;7(1):40–3.
160. Nosaka N, Fujiwara T, Knaup E, Okada A, Tsukahara H. Validity of mothers' reports of children's weight in Japan. Acta Med Okayama. 2016;70(4):255–9.
161. Sahar M, Nordin N, Nor A, Zakaria M. Weight estimation method in Emergency Department in Malaysia: is Broselow Tape (BT) reliable? Jurnal Teknologi (Sciences & Engineering). 2016;78(6–8):113–7.
162. So J, Chow E, Cattermole G, Rainer T. A comparison of the performance of different age-based paediatric weight estimation formulae in Hong Kong children. Hong Kong Journal of Emergency Medicine. 2016;23:3–12.
163. So JL, Chow EP, Cattermole GN, Graham CA, Rainer TH. Finger counting method is more accurate than age-based weight estimation formulae in estimating the weight of Hong Kong children presenting to the emergency department. Emergency Medicine Australasia. 2016;28(6):691–7.
164. Bowen L, Zyambo M, Snell D, Kinnear J, Bould MD. Evaluation of the accuracy of common weight estimation formulae in a Zambian paediatric surgical population. Anaesthesia. 2017;72(4):470–8.
165. Reilly JJ. Mid-upper arm circumference (MUAC): new applications for an old measure. Arch Dis Child. 2017;102(1):1–2.
166. Trainarongsakul T, Sanguanwit P, Rojcharoenchai S, Sawanyawisuth K, Sittichanbuncha Y. The RAMA Ped Card: does it work for actual weight estimation in child patients at the emergency department. World J Emerg Med. 2017;8(2):126–30.
167. Waseem M, Chen J, Leber M, Giambrone AE, Gerber LM. A reexamination of the accuracy of the Broselow Tape as an instrument for weight estimation. Pediatr Emerg Care. 2017. doi:10.1097/PEC.0000000000000982. [Epub ahead of print].

Supplementary Table 1 Data from each study included in the meta-analysis, including subgroup data where available. Imputed data is indicated by bold type.

Ali Formula							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
Trinidad 2012	1723	All	-3.1 **[16.9]**	**-36.2, 30.0**	47.5	**75.6**	Ali
Advanced Paediatric Life Support (APLS) formula (new)							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
USA 2012	2253	All	-6.8 **[21.2]**	-48.4, 34.8	36.3	66.7	Cantle
	1256	1-5yrs	-12.3 **[15.7]**	-43.2, 18.5	42.0	75.1	
	997	6-10yrs	0.2 **[24.9]**	-48.5, 49.0	29.2	56.1	
UK 2012	599	All	**-6.1 [36.7]**	**-78.1, 65.9**	**21.2**	**40.9**	Seddon
	184	<1yr	-3.9 [39.4]	**-81.1, 73.3**	**19.9**	**38.7**	
	275	1-5yrs	-9.4 [34.7]	**-77.4, 58.6**	**21.9**	**42.2**	
	102	6-10yrs	-0.4 [38.7]	**-76.3, 75.5**	**20.4**	**39.5**	
	38	11-12yrs	-8.0 [28.9]	**-64.6, 48.6**	**26.1**	**49.5**	
Australia 2013	28508	All			25.1		Graves
	294	<1yr			33.3		
	4425	1-5yrs			35.5		
	15246	6-10yrs			23.4		
	8543	11-14yrs			22.7		
New Zealand 2014	376	5-10yrs			39.1		Britnell
Mali 2014	345	All	23.6 [21.4]	-18.3, 65.5	15.0	28.1	Dicko
UK 2014	9436	All	**12.2 [26.2]**	**-39.2, 63.7**	**26.8**	**50.8**	Flannigan
	5204	Infants	**20.2 [24.5]**	**-27.8, 68.2**	**23.0**	**44.6**	
	2907	1-5yrs	**-2.9 [19.1]**	**-40.3, 34.6**	**39.5**	**69.9**	
	1325	6-12yrs	**13.6 [30.5]**	**-46.2, 73.5**	**23.3**	**44.8**	
Australia 2017	183	All			45.8		O'Leary
	18	<1yr			44.4		
	78	1-5yrs			35.9		
	68	6-10yrs			47.1		
	19	11-14yrs			62.9		
USA 2015	324	All	18.7		36.4	65.7	Chavez
	65	<1yr	15.8				
	163	1-5yrs	16.6				
	96	6-12yrs	21.2				
Nigeria 2016	300	All	3.1 [18.1]	**-32.3, 38.5**	**48.2**		Aliyu
Hong Kong 2016	4600	All	0.1	-39.8, 40.1	42.9	74.6	So
Zambia 2017	1289	All	15.4		29.6	54.0	Bowen
Thailand 2017	430	All	8.3 [26.3]	**-43.2, 59.8**	29.5	54.2	Samerchua
South Africa 2017	1066	All	6.1	-40.4, 52.7	40.8	69.8	Wells
Argall formula							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
Australia 2007	410	All			37.0		Nguyen
		<10kg			85.0		
Croatia 2010	156	All			46.0		Lulic
Hong Kong 2011	1248	All	5.3 **[20.1]**	-34.1, 44.7	34.6	80.6	Cattermole
	609	1-6yrs	3.8 **[17.3]**	-30.2, 37.8	40.7	83.6	
	639	7-10yrs	6.8 **[22.3]**	-37.0, 50.5	28.8	77.8	
	463	<20kg	10.3 **[17.1]**	-23.2, 43.7	36.5	82.9	
	503	20-30kg	11.7 **[17.2]**	-22.0, 45.4	29.2	77.7	
	282	>30kg	-14.1 **[17.1]**	-47.7, 19.5	41.1	81.9	
UK 2011	140314	All	-5.0 **[21.6]**	-47.4, 37.3	39.0	**63.3**	Marlow
USA 2012	1013	All	-4.6 [20.5]	**-44.8, 35.6**	19.8	34.4	Abdel-Rahman
UK 2012	377	All	**-5.8 [39.1]**	**-82.4, 70.9**	**20.0**	**38.7**	Seddon
	275	1-5yrs	-6.6 **[39.8]**	**-84.6, 71.4**	**19.6**	**38.0**	
	102	6-10yrs	-3.5 **[37.1]**	**-76.2, 69.2**	**21.2**	**40.9**	
India 2014	249	All	31.5 [31.0]	**-29.3, 92.3**	10.0	23.0	Batmanabane
Australia 2014	1739	All	3.4 [16.1]	**-28.2, 35.0**	46.5		Allison
Zambia 2017	952	All	18.1		25.7	50.1	Bowen
South Africa 2017	883	All	9.5	-34.9, 53.9	36.5	67.3	Wells
Australian Resuscitation Council (ARC) formula							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
Australia 2007	1572	All	**-12.2 [16.7]**	**-45.0, 20.5**	**35.6**	**65.3**	Thompson
	806	1-5yrs	-12.1 [13.3]	**-38.2, 14.0**	**38.9**	**71.6**	
	766	6-14yrs	-12.4 [19.7]	**-51.0, 26.2**	**32.4**	**60.0**	
Hong Kong 2011	1248	All	-10.1 **[19.8]**	-48.9, 28.7	42.1	83.8	Cattermole
	609	1-6yrs	-7.5 **[16.5]**	-39.8, 24.7	49.4	87.8	
	639	7-10yrs	-12.5 **[22.2]**	-56.1, 31.2	34.1	80.8	
	463	<20kg	0.1 **[13.2]**	-25.8, 25.9	57.0	90.7	
	503	20-30kg	-7.3 **[16.1]**	-38.9, 24.3	44.7	84.9	
	282	>30kg	-31.7 **[18.3]**	-67.6, 4.3	2.5	68.8	
USA 2012	1752	All	-16.9 [19.0]	-54.1, 20.3	27.5	51.0	Abdel-Rahman
Mali 2014	446	All	9.6 [18.0]	**-25.7, 44.8**	40.8	68.7	Dicko
India 2014	350	All	18.1 [27.3]	**-35.4, 71.6**	23.0	41.0	Batmanabane
Zambia 2017	1071	All	11.1		36.8	64.1	Bowen

South Africa 2017	966	All	-2.0 [19.8]	-40.9, 36.9	45.3	75.6	Wells
Best Guess (BG) Formula							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
Australia 2007	410	All				70.4	Kelly
	219	1-5yrs				76.0	
	191	5-11yrs				64.0	
Australia 2007	410	All			42.1		Krieser
Australia 2007	1843	All	5.0 [21.2]	-36.6, 46.5	42.3	72.4	Thompson
	271	<1yr	4.2 [23.8]	-42.4, 50.8	43.8	76.1	
	806	1-5yrs	2.6 [15.6]	-28.0, 33.2	53.1	83.4	
	766	6-14yrs	7.7 [24.7]	-40.7, 56.1	30.5	59.5	
Australia 2010	1426	All	3.6 [19.6]	-34.8, 41.9	38.4	65.1	Casey
	174	<1yr	6.3 [21.6]	-36, 50	34.3	76.0	
	520	1-4yrs	1.7 [15.2]	-28, 32	48.7	83.0	
	541	5-10yrs	4.5 [20.2]	-35, 45	37.1	67.0	
	191	11-14yrs	3.4 [25.2]	-46, -54	30.6	59.0	
Croatia 2010	209				46.0		Lulic
Hong Kong 2011	1248	All	10.1 [20.0]	-29.1, 49.4	31.9	83.0	Cattermole
	609	1-6yrs	6.2 [16.8]	-26.8, 39.2	39.4	84.1	
	639	7-10yrs	13.9 [22.0]	-29.3, 57.0	24.7	82.0	
	463	<20kg	14.1 [14.9]	-15.2, 43.3	57.0	90.7	
	503	20-30kg	15.6 [20.3]	-24.1, 55.3	36.9	83.8	
	282	>30kg	-6.1 [18.4]	-42.1, 29.8	24.7	82.0	
South Africa 2011	2832	All	15.4 [22.0]	-27.7, 58.5	31.5	52.9	Geduld
	1710	1-5yrs	12.7 [19.1]	-24.7, 50.1	40.5	60.5	
	1122	6-10yrs	25.0 [24.0]	-22.0, 72.0	17.7	38.7	
UK 2011	140314	All	5.6 [22.3]	-38.0, 49.3	43.7	61.5	Marlow
USA 2012	1563	All	1.0 [21.7]	-41.5, 43.5	31.2	54.0	Abdel-Rahman
Korea 2012	124095	All	5.0 [18.7]	-32.1, 41.6	39.4	69.8	Park
	19854	Infant	2.3 [23.0]	-42.9, 47.4	33.5	61.3	
	40612	Preschool	4.2 [14.0]	-23.3, 31.6	50.5	82.9	
	63629	School	6.3 [19.6]	-32.1, 44.7	37.2	66.8	
UK 2012	599	All	3.6 [35.8]	-66.6, 73.9	21.9	42.2	Seddon
	184	<1yr	3.4 [37.4]	-69.9, 76.7	21.0	40.6	
	275	1-5yrs	4.7 [34.0]	-61.9, 71.3	22.9	44.0	
	102	6-10yrs	1.8 [39.7]	-76.0, 79.6	19.9	38.5	
	38	11-12yrs	2.1 [29.2]	-55.1, 59.3	26.7	50.5	
Australia 2013	37038	All			24.6		Graves
	219	<1yr			29.7		
	4425	1-5yrs			36.2		
	15246	6-10yrs			22.4		
	17148	11-14yrs			23.5		
India 2014	347	All	41.3 [33.2]	-23.8, 106.4	10.0	24.0	Batmanabane
Australia 2014	2102	All	16.4 [15.2]	-13.4, 46.2	48.9		Allison
	363	<1yr	4.4 [16.6]	-25.2, 27.8	45.3		
	1739	1-5yrs	19.0 [13.5]	-7.5, 45.5	54.1		
Nigeria 2014	2754	All	21.7 [19.3]	-16.1, 59.5	22.0		Omisanjo
Australia 2015	199	All			47.8		O'Leary
	18	<1yr			61.1		
	78	1-5yrs			51.3		
	68	6-10yrs			47.1		
	35	11-14yrs			31.4		
Hong Kong 2016	4600	All	10.9 [20.6]	-29.4, 51.3	38.5	68.2	So
Zambia 2017	1381	All	30.4		16.4	34.3	Bowen
South Africa 2017	1075	All	16.1	-32.0, 64.2	32.3	59.2	Wells
Chinese Age Weight Rule 1 (CAWR-1)							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
Hong Kong 2011	1248	All	0.8 [20.5]	-39.3, 40.9	36.1	80.0	Cattermole
	609	1-6yrs	-2.0 [18.0]	-37.2, 33.1	42.4	81.9	
	639	7-10yrs	3.5 [22.2]	-40.1, 47.2	30.2	78.2	
	463	<20kg	4.1 [18.4]	-31.9, 40.2	40.0	81.9	
	503	20-30kg	7.9 [17.9]	-27.1, 42.9	33.2	78.7	
	282	>30kg	-17.2 [17.2]	-51.0, 16.6	35.1	79.4	
USA 2012	2253	All	-12.5 [22.3]	-56.1, 31.2	30.9	58.8	Cantle
	1256	1-5yrs	-17.2 [18.9]	-54.2, 19.7	28.5	57.4	
	997	6-10yrs	-6.5 [24.7]	-54.8, 41.9	34.0	60.6	
Hong Kong 2016	4600	All	-5.2 [20.7]	-45.8, 35.3	35.3	65.6	So
Zambia 2017	952	All	10.5		31.0	54.4	Bowen
South Africa 2017	963	All	3.9 [22.6]	-40.4, 48.2	39.9	69.6	Wells
Chinese Age Weight Rule 2 (CAWR-2)							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
Hong Kong 2011	1248	All	1.6 [19.6]	-36.8, 40.0	38.7	83.3	Cattermole
	609	1-6yrs	2.1 [16.6]	-30.5, 34.7	44.7	85.4	
	639	7-10yrs	1.2 [22.1]	-42.1, 44.4	30.2	78.2	
	463	<20kg	5.8 [14.1]	-21.8, 33.5	49.5	84.0	
	503	20-30kg	2.9 [17.6]	-31.6, 37.4	41.7	84.7	
	282	>30kg	-20.5 [18.0]	-55.7, 14.7	25.9	76.6	
USA 2012	2253	All	-6.0 [20.2]	-45.7, 33.6	44.7	74.6	Cantle
	1256	1-5yrs	-3.9 [15.8]	-34.8, 27.1	51.3	84.2	
	997	6-10yrs	-8.8 [24.5]	-56.8, 39.2	36.3	62.6	
South Africa 2017	883	All	7.9 [21.0]	-33.3, 49.1	40.7	70.8	Wells

European Paediatric Life Support (EPLS) formula (original APLS)							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
UK 1999	50	All			34.0		Dearlove
Australia 2002	254	All	-13.9 [16.4]	-46.6, 18.8	33.4	62.6	Black
	132	10-25kg	-4.7 [11.7]	-27.6, 18.2	57.0	88.7	
	102	25-40kg	-20.0 [12.9]	-45.2, 5.2	20.9	49.9	
	21	>40kg	-42.0 [9.4]	-60.4,-23.6	0.0	1.0	
UK / New Zealand 2005	237	All	-21.2 [24.8]	-69.8, 27.4	22.2	43.2	Luscombe
New Zealand 2005	909	All	-7.3 [30.2]	-66.5, 52.0	25.2	48.0	Theron
	226	Maori	-5.5 [33.9]	-71.9, 60.9	22.9	44.0	
	420	Pacific	-11.1 [26.6]	-63.2, 41.0	27.0	51.0	
	160	European	-5.9 [25.7]	-56.2, 44.4	29.5	55.2	
	79	Asian	5.1 [40.1]	-73.5, 83.7	19.5	37.9	
Australia 2007	410	All			34.0		Krieser
UK 2007	13988	All	-18.8 [22.9]	-63.6, 26.1	24.6	47.6	Luscombe
	9010	1-5yrs	-13.9 [17.9]	-48.9, 21.1	32.3	60.4	
	4978	6-10yrs	-27.9 [27.6]	-82.1, 26.2	17.3	34.6	
Australia 2007	1572	All	-15.9 [17.3]	-49.1, 17.3	29.9	57.5	Thompson
	806	1-5yrs	-12.1 [13.3]	-38.2, 14.0	38.9	71.6	
	766	6-14yrs	-19.9 [19.3]	-57.7, 17.9	24.3	48.3	
Malawi 2007	100	All	10.6 [16.8]	-22.3, 43.5	37.6	67.8	Pollock
England 2009	544	All	-12.0 [17.3]	-46.0, 22.0	35.2	64.6	Sandell
Hong Kong 2010	1370	All	-13.4 [20.7]	-54.0, 27.2	39.0	83.1	Cattermole
	449	1-5yrs	-7.4 [16.1]	-39.0, 24.1	49.2	87.5	
	473	6-8yrs	-10.6 [20.1]	-49.9, 28.7	42.9	84.4	
	448	9-11yrs	-22.2 [22.6]	-66.4, 22.0	24.6	77.2	
Australia 2010	1061	All	-15.0 [14.6]	-43.6, 13.5	32.3	64.4	Casey
	520	1-4yrs	-12.6 [13.0]	-38, 16	38.0	71.0	
	541	5-10yrs	-17.4 [15.6]	-48, 18	27.8	58.0	
Croatia 2010	156	All			36.0		Lulic
Hong Kong 2011	1248	All	-11.7 [20.2]	-51.2, 27.7	41.6	84.2	Cattermole
	609	1-6yrs	-7.5 [16.5]	-39.8, 24.7	49.4	87.8	
	639	7-10yrs	-15.7 [22.4]	-59.6, 28.2	34.1	80.8	
	463	<20kg	0.1 [13.2]	-25.8, 25.9	57.0	90.7	
	503	20-30kg	-8.5 [14.4]	-36.8, 19.9	49.3	86.9	
	282	>30kg	-36.9 [15.9]	-68.1, -5.7	2.5	68.8	
South Africa 2011	2832	All	-3.3 [17.1]	-36.8, 30.2	57.9	74.9	Geduld
	1710	1-5yrs	-2.4 [16.5]	-34.7, 29.9	58.4	77.0	
	1122	6-10yrs	-4.6 [17.9]	-39.7, 30.5	57.1	72.1	
Australia 2011	410	All			34.0		Kelly
UK 2011	64197	All	-22.0 [23.8]	-68.6, 24.6	21.8	42.8	Luscombe
	36273	1-5yrs	-16.8 [17.4]	-50.9, 17.3	28.6	55.6	
	27924	6-10yrs	-28.7 [28.8]	-85.1, 27.7	16.9	33.6	
UK 2011	140314	All	-13.1 [18.0]	-48.4, 22.1	35.1	61.6	Marlow
USA 2012	1013	All	-14.9 [17.2]	-48.6, 18.8	17.8	33.2	Abdel-Rahman
Trinidad 2012	1723	All	-5.8 [16.9]	-38.9, 27.3	45.6	73.6	Ali
USA 2012	2253	All	-17.8 [21.1]	-59.1, 23.5	34.8	64.4	Cantle
	1256	1-5yrs	-12.3 [15.7]	-43.2, 18.5	42.0	75.1	
	997	6-10yrs	-24.6 [24.6]	-72.9, 23.7	25.7	51.1	
Singapore 2012	875	All	-7.6 [17.4]	-41.7, 26.5	45.7	70.6	Loo
Korea 2012	104241	All	-11.5 [14.0]	-38.9, 16.0	39.5	71.6	Park
	40612	Preschool	-9.5 [11.9]	-32,9, 13.6	46.6	80.5	
	63629	School	-13.2 [15.4]	-43.3, 17.0	35.2	65.5	
UK 2012	377	All	-12.8 [36.1]	-83.4, 57.9	20.5	39.7	Seddon
	275	1-5yrs	-9.4 [33.8]	-75.6, 56.8	22.4	43.1	
	102	6-10yrs	-21.8 [40.2]	-100.6, 57.0	17.0	33.3	
Botswana 2012	531	All			37.3		Wozniak
Australia 2013	19671	All			29.0		Graves
	4425	1-5yrs			35.5		
	15246	6-10yrs			27.2		
Kenya 2013	967	All	5.2 [19.0]	-32.0, 42.4	38.8	69.0	House
Egypt 2013	394	All	-0.6 [24.3]	-48.2, 47.0	37.9	70.6	Hegazy
USA 2013	207	All			33.0	72.0	Young
India 2014	249	All	13.9 [24.8]	-34.7, 62.5	17.0	45.0	Batmanabane
Australia 2014	1739	All	-0.1 [12.0]	-23.6, 23.4	59.5		Allison
New Zealand 2015	305	All			46.2		Britnell
Ireland 2015	3155	All	-20.3 [17.2]	-13.4, 54.0	23.6	48.3	Skrobo
Japan 2016	237	All			63.8	89.9	Nosaka
Hong Kong 2016	4600	All	-8.3 [17.0]	-41.7, 25.0	44.6	76.0	So
Zambia 2017	1071	All	8.1		38.0	66.9	Bowen
Thailand 2017	274	All	-4.4 [24.4]		33.9	65.7	Samerchua
South Africa 2017	963	All	-4.2 [19.2]	-41.9, 33.5	46.8	75.8	Wells

Garwood Formula							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
Egypt 2013	394	1-10yrs	14.4 [28.4]	-41.3, 70.1	35.8	60.9	Hegazy
Zambia 2017	1071	All	20.0		24.4	48.0	Bowen
South Africa 2017	996	All	9.2 [23.3]	-36.5, 54.8	35.3	65.9	Wells

Leffler formula							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
USA 1997	117	All	-13.6 [17.5]	-47.9, 20.7	46.2	61.5	Leffler
UK / New Zealand 2005	237	All	-9.1 [23.4]	-55.0, 36.7	30.8	57.2	Luscombe

New Zealand 2005	909	All	-7.3 [20.3]	-47.1, 32.6	35.6	64.5	Theron
	226	Maori	-6.4 [25.6]	-56.6, 43.8	29.5	55.1	
	420	Pacific	-11.0 [17.1]	-44.5, 22.5	36.7	66.6	
	160	European	-2.0 [18.2]	-37.7, 33.7	41.5	72.5	
	79	Asian	-0.5 [18.9]	-37.5, 36.5	40.3	71.0	
USA 2009	1011	All	11.0 [30.5]	-48.7, 70.7	24.1	46.1	So
	471	<10kg	24.0 [33.0]	-40.7, 88.7	18.4	36.1	
	382	10-25kg	9.0 [17.0]	-24.3, 42.3	39.2	69.7	
	119	25-40kg	-16.0 [14.0]	-43.4, 11.4	30.2	60.7	
	39	>40kg	-44.0 [12.0]	-67.5, 20.5	0.2	2.3	
	166	UW	35.0 [40.0]	-43.4, 113.0	13.6	26.9	
	650	NW	11.0 [23.0]	-34.1, 56.1	30.2	56.3	
	195	OW	-11.0 [28.0]	-65.9, 43.9	25.9	49.2	
Hong Kong 2011	1248	All	-2.1 [21.0]	-43.2, 39.0	37.6	83.5	Cattermole
	609	1-6yrs	4.2 [16.8]	-28.8, 37.1	41.7	85.9	
	639	7-10yrs	-8.1 [22.7]	-52.5, 36.4	33.6	81.2	
	463	<20kg	12.2 [12.5]	-12.3, 36.7	36.9	84.9	
	503	20-30kg	0.2 [13.6]	-26.5, 26.9	53.7	89.7	
	282	>30kg	-29.6 [16.0]	-60.9, 1.7	9.9	70.2	
UK 2011	140314	All	-1.4 [21.4]	-43.3, 40.5	44.7	64.9	Marlow
USA 2012	1199	All	-5.5 [19.6]	-43.9, 32.9	24.8	42.9	Abdel-Rahman
Korea 2012	124095	All	-1.7 [17.3]	-35.6, 32.1	43.5	75.0	Park
	19854	Infant	-6.4 [19.7]	-45.1, 32.2	37.0	66.5	
	40612	Preschool	4.2 [14.0]	-23.3, 31.6	50.5	82.9	
	63629	School	-4.9 [17.3]	-38.7, 28.9	42.1	73.4	
India 2014	247	All	27.8 [28.5]	-28.1, 83.7	11.0	23.0	Batmanabane
Zambia 2017	1024	All	22.4		22.5	44.2	Bowen
South Africa 2017	972	All	7.7 [21.9]	-35.3, 50.7	37.6	68.3	Wells

Luscombe formula

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
UK 2007	13988	All	-2.5 [18.7]	-38.8, 34.5	40.4	71.1	Luscombe
	9010	1-5yrs	-4.5 [17.3]	-38.3, 29.4	42.4	73.7	
	4978	6-10yrs	1.1 [21.1]	-40.2, 42.4	36.4	65.6	
Malawi 2007	100	All	17.0 [17.0]	-16.3, 50.3	28.4	55.5	Pollock
Croatia 2010	156	All			46.0		Lulic
Hong Kong 2011	1248	All	9.6 [19.9]	-29.4, 48.6	30.6	79.8	Cattermole
	609	1-6yrs	9.3 [16.9]	-23.9, 42.4	35.3	71.3	
	639	7-10yrs	9.9 [22.3]	-33.9, 53.8	26.1	78.4	
	463	<20kg	16.0 [16.0]	-15.3, 47.2	28.3	83.2	
	503	20-30kg	15.3 [16.6]	-17.2, 47.9	25.6	76.1	
	282	>30kg	-11.1 [17.0]	-44.5, 22.3	43.3	80.9	
South Africa 2011	2832	All	12.4 [22.6]	-31.9, 56.7	34.2	55.6	Geduld
	1710	1 to 5yrs	8.6 [20.6]	-31.8, 49.0	44.4	62.7	
	1122	6 to 10yrs	22.6 [22.8]	-22.1, 67.3	18.5	42.9	
Australia 2011	410	All			45.4		Kelly
UK 2011	64197	All	-2.1 [20.1]	-41.5, 37.4	37.9	67.8	Luscombe
	36273	1-5yrs	-3.8 [16.5]	-36.2, 28.6	44.5	76.2	
	27924	6-10yrs	0.2 [23.8]	-46.4, 46.8	32.6	59.9	
UK 2011	140314	All	0.9 [21.9]	-42.1, 43.8	43.1	63.8	Marlow
USA 2012	1013	All	0.8 [20.8]	-40.0, 41.6	20.2	35.9	Abdel-Rahman
Trinidad 2012	1723	All	5.0 [19.1]	-32.4, 42.4	42.3	63.8	Ali
Singapore 2012	875	All	7.4 [22.2]	-36.2, 51.0	37.7	60.6	Loo
UK 2012	652	All	-0.5 [32.9]	-64.9, 64.0	23.9	45.7	Seddon
	377	1-5yrs	-0.5 [31.3]	-61.8, 60.8	25.1	47.7	
	275	6-10yrs	-0.4 [34.9]	-68.8, 68.0	22.6	43.3	
Botswana 2012	531	All			7.2		Wozniak
Australia 2013	36819	All			25.5		Graves
	4425	1-5yrs			33.7		
	15246	6-10yrs			23.4		
	17148	11-14yrs			24.0		
USA 2013	207	All			52.0	84.0	Young
India 2014	249	All	38.0 [31.5]	-23.7, 99.7	6.0	14.0	Batmanabane
Australia 2014	1739	All	11.1 [16.0]	-20.3, 42.5	46.8		Allison
Australia 2015	181	All			47.3		O'Leary
	78	1-5yrs			38.5		
	68	6-10yrs			47.1		
	35	11-14yrs			37.1		
Ireland 2015	3155	All	3.3 [20.1]	-36.1, 42.7	37.6	67.4	Skrobo
Hong Kong 2016	4600	All	7.0 [20.7]	-33.7, 47.6	39.8	69.7	So
Zambia 2017	952	All	26.5		20.0	38.7	Bowen
South Africa 2017	883	All	15.5 [23.2]	-30.0, 60.9	29.1	59.1	Wells

Nelson formula

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
Hong Kong 2011	1248	All	-7.9 [19.5]	-46.1, 30.3	42.4	84.2	Cattermole
	609	1-6yrs	-7.5 [16.5]	-39.8, 24.7	49.4	87.8	
	639	7-10yrs	-8.2 [22.0]	-51.3, 34.9	35.7	80.8	
	463	<20kg	0.2 [13.4]	-26.0, 26.4	57.0	90.1	
	503	20-30kg	-4.4 [17.8]	-39.2, 30.4	42.1	82.5	
	282	>30kg	-27.3 [17.9]	-62.3, 7.7	18.8	75.9	
USA 2012	1399	All	-7.7 [19.5]	-45.9, 30.5	25.9	49.4	Abdel-Rahman
Korea 2012	124095	All	-10.0 [14.2]	-37.8, 17.9	42.1	74.2	Park
	19854	Infant	-8.3 [12.7]	-33.2, 16.5	47.8	80.9	
	40612	Preschool	-9.5 [11.9]	-32,9, 13.8	46.6	80.5	

	63629	School	-10.6 [15.9]	**-41.8, 20.5**	**38.7**	**69.6**	
Kenya 2013	967	All	10.4 **[19.0]**	**-26.8, 47.6**	**35.0**	**63.9**	House
Mali 2014	369	All	19.8 [21.6]	**-22.5, 62.1**	23.5	40.4	Dicko
Nigeria 2014	2754	All	6.4 [15.1]	**-23.2, 35.9**	45.5	77.6	Omisanjo
India 2014	329	All	28.2 [31.6]	**-33.7, 90.1**	18.0	31.0	Batmanabane
Australia 2014	2102	All	-1.2 [13.3]	**-27.3, 24.9**	54.8		Allison
	363	<1yr	1.3 [13.5]	**-25.2, 27.8**	54.3		
	1739	1-5yrs	-0.1 [12.0]	**-23.6, 23.4**	59.5		
Zambia 2017	1024	All	11.1		36.7	63.7	Bowen
South Africa 2017	1052	All	4.0 [21.7]	-38.6, 46.6	44.2	73.0	Wells
Park Formula							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
Korea 2012	124095	All	0.9 [18.2]	**-34.9, 36.6**	**41.7**	**72.8**	Park
	19854	Infant	2.3 [23.0]	**-42.9, 47.4**	**33.5**	**61.3**	
	40612	Preschool	-2.7 [12.9]	**-28.0, 22.6**	**55.2**	**87.1**	
	63629	School	2.7 [19.1]	**-34.7, 40.1**	**39.6**	**70.0**	
Japan 2016	237	All			45.1	75.5	Nosaka
Zambia 2017	1381	All	25.0		21.9	42.9	Bowen
South Africa 2017	1075	All	10.9 [23.9]	-36.0, 57.8	39.0	66.0	Wells
Shann formula							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
New Zealand 2005	909	All	**-0.8 [33.9]**	**-67.2, 65.6**	**23.2**	**44.5**	Theron
	226	Maori	-2.3 [37.9]	**-76.6, 72.0**	**20.8**	**40.2**	
	420	Pacific	-3.7 [29.9]	**-62.3, 54.9**	**26.0**	**49.3**	
	160	European	1.8 [28.9]	**-54.8, 58.4**	**27.0**	**51.0**	
	79	Asian	13.9 [44.9]	**-74.1, 101.9**	**16.8**	**32.9**	
Hong Kong 2011	1248	All	-6.5 **[20.3]**	-46.3, 33.4	42.7	83.7	Cattermole
	609	1-6yrs	-1.5 **[16.6]**	-34.1, 31.1	50.7	87.4	
	639	7-10yrs	-11.2 **[22.3]**	-55.0, 32.7	35.1	80.3	
	463	<20kg	6.3 **[12.8]**	-18.7, 31.4	53.6	88.1	
	503	20-30kg	-3.8 **[14.3]**	-31.8, 24.2	53.3	87.7	
	282	>30kg	-32.1 **[16.2]**	-63.8, -0.5	6.0	69.5	
USA 2012	1752	All	-17.4 [20.9]	**-58.4, 23.7**	26.3	48.6	Abdel-Rahman
Korea 2012	104241	All	-8.6 [16.4]	**-40.7, 23.5**	**40.6**	**71.6**	Park
	40612	Preschool	-2.7 [12.9]	**-27.9, 22.6**	**55.2**	**87.1**	
	63629	School	-12.4 [17.2]	**-46.1, 21.3**	**34.8**	**64.1**	
Egypt 2013	394	1-10yrs	-5.3 **[26.3]**	**-56.8, 46.2**	39.8	65.7	Hegazy
India 2014	350	All	18.4 [27.1]	**-34.7, 71.5**	22.0	42.0	Batmanabane
New Zealand 2015	305	All			53.1		Britnell
Zambia 2017	1071	All	16.1		30.7	56.4	Bowen
South Africa 2017	996	All	1.5 [21.0]	-39.7, 42.7	43.8	71.3	Wells
Theron formula							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
USA 2009	1011	All	**56.0 [77.4]**	**-95.9, 207.6**	**7.9**	**15.8**	So
	471	<10kg	106.0 [85.0]	**-60.6, 273**	**4.3**	**8.7**	
	382	10-25kg	12.0 [26.0]	**-40.0, 63.0**	**27.1**	**51.2**	
	119	25-40kg	20.0 [34.0]	**-46.6, 86.6**	**19.6**	**38.0**	
	39	>40kg	-10.0 [23.0]	**-55.1, 35.1**	**30.8**	**57.2**	
	166	UW	101.0 [105.0]	**-105.0, 307.0**	**4.8**	**9.6**	
	650	NW	56.0 [69.0]	**-79.2, 191.0**	**8.3**	**16.6**	
	195	OW	15.0 [48.0]	**-79.1, 109.0**	**15.7**	**30.9**	
Hong Kong 2011	1248	All	19.7 **[23.1]**	-25.5, 65.0	24.8	84.7	Cattermole
	609	1-6yrs	8.2 **[18.0]**	-27.0, 43.3	36.8	83.6	
	639	7-10yrs	30.7 **[22.1]**	-12.6, 74.0	13.5	85.8	
	463	<20kg	14.7 **[25.5]**	-22.1, 51.6	33.0	87.3	
	503	20-30kg	27.9 **[25.1]**	-21.2, 77.0	16.1	84.7	
	282	>30kg	13.3 **[21.5]**	-28.9, 55.4	27.0	80.5	
USA 2012	1013	All	9.6 [26.4]	**-42.1, 61.3**	19.6	33.4	Abdel-Rahman
Botswana 2012	531	All			8.5		Wozniak
India 2014	249	All	51.4 [42.6]	**-32.1, 134.9**	7.0	14.0	Batmanabane
New Zealand 2015	305	All			27.5		Britnell
Zambia 2017	920	All	27.0		19.9	37.3	Bowen
South Africa 2017	1085	All	29.2 [37.3]	-43.9, 102.4	25.4	45.4	Wells
Guesses by healthcare providers (HCPs)							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
USA 1990	138	All					Hughes
England 1997	75	All	**1.3 [34.4]**	**-66.1, 68.7**	**22.9**	**43.9**	Greig
USA 1999	100	All	-5.9 **[19.2]**	-43.5, 34.6	**38.1**	**68.0**	Harris
Malawi 1999	142	All				54.0	Molyneux
Japan 2002	48	All				**57.5**	Uesegi
USA 2009	704	All			56.0		Partridge
Australia 2010	468	All			**17.1**	**45.7**	Williams
USA 2011	372	All	-2.8 [21.8]	**-45.5, 39.9**	44.0	63.7	Rosenberg
USA 2013	179	All				82.4	Lim
Estimates by parents							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
USA 1997	117	All	-6.8 [9.8]	-26.0, 12.4	80.3	**90.8**	Leffler
UK 1999	50	All			25.0		Dearlove

Israel 1999	233	All			73.4	87.5	Goldman
	138	< 5yrs	0.1 [18.5]	-36.2, 36.3	41.1	72.0	
USA 1999	100	All	-1.3 [19.5]	-38.9, 37.7	39.1	69.4	Harris
Australia 2007	410	All			78.0		Krieser
USA 2009	777	All			79.0		Partridge
Belgium 2011	71	All					Huybrechts
Thailand 2012	595	All			85.2		Trakulsrichai
	111	<10kg			80.2		
	372	20-25kg			87.1		
	104	25-40kg			84.6		
	8	>40kg			75.0		
USA 2013	207	All			54.0	79.0	Young
Japan 2016	237	All			94.9	98.7	Nosaka
Thailand 2017	344	All	-2.6 [7.1]	-16.5, 11.3	88.7	97.4	Samerchua

Broselow Tape

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
USA 1988 BT 1998	937	All			59.7	89.6	Lubitz
	395	3.5-10kg			55.9	85.6	
	449	10-25kg			65.0	95.1	
	93	>25kg			49.5	83.9	
UK 1999 *BT version NR*	50	All			60.0		Dearlove
Hong Kong 2000 *BT version NR*	909	All			69.5	91.3	Kun
	129	<10kg			56.6	83.7	
	631	10-25kg			76.2	95.2	
	149	>25kg			52.3	81.2	
Australia 2002 *BT version NR*	339	All	-2.0 [11.7]	-24.9, 20.9	60.0	90.8	Black
	121	<10kg	-0.6 [12.3]	-24.7, 23.5	58.3	89.6	
	132	10-25kg	-0.4 [9.4]	-18.8, 20.9	71.2	96.6	
	86	25-40kg	-6.4 [12.8]	-31.5, 18.7	51.1	83.6	
Switzerland 2002 *BT version NR*	585	All			65.0		Hofer
	339	≤20kg			66.0		
	246	>20kg			65.0		
New Zealand 2005 *BT version NR*	909	All	-9.5 [12.6]	-34.1, 15.1	45.5	78.8	Theron
	226	Maori	-11.1 [13.6]	-37.9, 15.6	40.7	73.2	
	420	Pacific	-11.0 [11.0]	-32.6, 10.5	43.6	79.1	
	160	European	-5.9 [12.5]	-30.4, 18.7	52.7	85.1	
	79	Asian	-3.9 [14.0]	-31.2, 23.5	50.8	83.1	
USA 2006 BT1998	7813	All	-3.9 [11.9]	-27.2, 19.4	55.3	86.6	Nieman
	2305	<10kg	-3.6 [12.7]	-28.4, 21.3	57.0	87.0	
	2660	10-20kg	-5.0 [10.9]	-26.4, 16.5	59.8	90.5	
	2848	>20kg	-3.8 [15.6]	-34.4, 26.7	53.2	78.7	
USA 2006 BT2005A	7671	All	-5.6 [11.9]	-28.9, 17.7	60.0	93.7	Nieman
	2305	<10kg	-2.9 [11.9]	-26.3, 20.5	62.3	89.7	
	2660	10-20kg	-5.7 [9.5]	-24.4, 13.0	62.0	93.0	
	2706	>20kg	-7.8 [13.4]	-34.0, 18.5	58.5	80.0	
USA 2007 *BT version NR*	300	All	-8.2 [19.1]	-45.6, 29.3	36.7	66.2	DuBois
	100	<10kg	-9.9 [15.5]	-40.3, 20.5	40.3	71.6	
	100	10-20kg	-7.1 [11.3]	-29.3, 15.0	53.6	86.5	
	100	20-36kg	-7.5 [26.9]	-60.1, 45.1	27.9	52.6	
USA 2007 BT2002B	1207	All				66.2	Hashikawa
	544	Birth				100	
	520	4-6yrs				70.0	
	143	10-12yrs				40.6	
Korea 2007 *BT version NR*	665	All			57.9		Jang
		≤26kg			64.5		
		>26kg			38.5		
Australia 2007 *BT version NR*	410	All			61.0		Krieser
India 2008 *BT version NR*	548	All	9.0 [13.1]	-16.8, 34.7	43.6	78.6	Ramarajan
	175	<10kg	2.4 [12.6]	-22.3, 27.1	52.6	88.1	
	197	10-18kg	11.3 [10.9]	-10.1, 32.7	44.7	78.6	
	176	>18kg	12.9 [13.5]	-13.6, 39.4	33.5	69.3	
Ireland 2009 *BT version NR*	545	All	-8.6 [13.2]	-34.5, 17.2	69.7	86.6	Anstett
India 2009 *BT version NR*	15000	All	-11.6		63.0		Cattamanchi
	4256	<10kg	-2.4				
	6076	10-18kg	-13.3				
	4668	>18kg	-17.9		57.0		
USA 2009 *BT version NR*	1011	All	-4.2 [21.6]	-46.4, 38.1	35.0	63.6	So
	471	<10kg	0.2 [24.0]	-46.8, 47.2	32.3	59.5	
	382	10-25kg	-3.7 [16.0]	-35.1, 27.7	45.7	77.7	
	119	25-40kg	-12.0 [17.0]	-45.3, 21.3	35.5	65.1	
	39	>40kg	-38.0 [12.0]	-61.5, -14.5	1.0	6.7	
	166	UW	24.0 [29.0]	-32.8, 80.8	19.4	38.1	
	650	NW	-3.0 [10.0]	-22.6, 16.6	66.1	94.5	
	195	OW	-28.0 [12.0]	-51.5, 43.9	6.6	25.2	
Australia 2009 *BT version NR*	475	All	-5.5 [12.9]	-31.2, 20.4	62.0	90.0	Stewart
	71	<10kg			68.0	88.0	
	316	10-25kg			63.0	93.0	
	88	>25kg			53.0	81.0	
Australia 2010 *BT version NR*	1235	All	-6.3 [14.6]	-34.8, 22.3	46.8	88.1	Casey
	174	<1yr	-7.2 [15.0]	-36, 22	44.8	83.0	
	520	1-4yrs	-7.0 [11.7]	-30, 16	52.8	91.0	
	541	5-10yrs	-5.3 [16.7]	-38, 58	43.1	87.0	

Hong Kong 2010 BT1998	1199 449 471 279	All 1-5yrs 6-8yrs 9-11yrs	-1.9 **[14.1]** -1.3 **[13.2]** -1.9 **[14.2]** -2.9 **[15.3]**	-29.5, 25.7 -27.2, 24.7 -29.7, 25.9 -32.8, 27.1	58.4 65.3 56.7 56.3	89.0 91.5 90.1 83.0	Cattermole
Croatia 2010 *BT version NR*	151	All			58.0		Lulic
Canada 2011 *BT version NR*	243	All	-11.9 **[14.9]**	-41.1, 17.3	48.1	81.5	Bourdeau
South Africa 2011 *BT version NR*	2832 1710 1122	All 1-5yrs 6-10yrs	-0.9 [14.3] -2.4 [13.8] 1.6 [14.7]	**-28.9, 27.1** **-29.4, 24.6** **-27.2, 30.4**	64.2 63.5 65.4	**83.7** **84.7** **82.4**	Geduld
USA 2011 BT2007B	657	All				46.9	Knight
USA 2011 BT2007B	372	All	-3.0 [15.3]	**-33.0, 27.0**	63.0	**80.0**	Rosenberg
USA 2012 *BT version NR*	1214	All	-6.1 [12.3]	**-30.1, 18.0**	37.4	53.9	Abdel-Rahman
Singapore 2012 BT2007B	875	All	-0.6 **[14.0]**	-28.1, 26.8	58.9	**72.3**	Loo
Canada 2012 BT2002A	6361	All	-7.1 [16.9]	-40.2, 26.0	56.3	84.5	Milne
Korea 2012 *BT version NR*	124095 19854 40612 63629	All Infant Preschool School	-4.8 [10.9] -4.9 [12.2] -4.0 [9.0] -5.5 [11.7]	**-26.1, 16.5** **-28.8, 19.0** **-21.6, 13.6** **-28.4, 17.4**	**59.6** **55.1** **68.8** **55.7**	**90.7** **87.1** **95.8** **87.8**	Park
USA 2012 *BT version NR*	118	All	-8.4 **[8.3]**	**-24.7, 7.9**	50.0	**91.9**	Sinha
Thailand 2012 *BT version NR*	595 111 372 104 8	All <10kg 10-25kg 25-40kg >40kg			56.1 60.4 59.7 43.3 0.0		Trakulsrichai
Botswana 2012 *BT version NR*	758	All				55.0	Wozniak
USA 2013 *BT version NR*	415	All	-4.1 [11.9]	**-27.4, 19.2**	58.6	90.8	Abdel-Rahman
Iran 2013 *BT version NR*	403	All			72.5		Akabarian
Sudan 2016 *BT version NR*	583 227 141 215	All Well nourish Malnourish Sev malnour				**15.1** 28.2 8.5 6.7	Clark
Australia 2013 BT2007B	24426 294 4421 14028 5683	All <1yr 1-5yrs 6-10yrs 11-14yrs			49.4 47.3 55.7 49.7 43.8		Graves
Kenya 2013 BT2007B	967	BT 2007B	2.2 **[12.7]**	**-22.7, 27.1**	**56.2**	**87.9**	House
Korea 2013	105072 104998 102316	BT 2005 A BT 2007 B BT 2011 A	-4.5 -3.9 -0.5		64.8 65.6 67.7		Suh
South Africa 2013 BT2007B	453 120 193 140	All <12kg 12-20kg >20kg	-3.8 [11.2] -2.4 [10.5] -3.0 [10.2] -6.0 [12.6]	**-25.8, 18.2** **-23.0, 18.2** **-23.0, 17.0** **-30.7, 18.7**	63.6 65.8 65.8 58.6	91.6 93.3 93.8 87.1	Wells
USA 2013 *BT version NR*	207	All			56.0	81.0	Young
Mali 2014 *BT version NR*	365	All	8.2 [10.4]	**-12.2, 28.6**	41.2	67.4	Dicko
India 2014 *BT version NR*	321	All	10.8 [16.3]	**-21.1, 42.7**	28.0	60.0	Batmanabane
Thailand 2014 BT2007A	3869	All	-3.6 [12.9]	**-28.9, 12.9**	58.0		Chiengkriwate
Australia 2014 BT2007B	2102 363 1739	All <1yr 1-5yrs	-0.4 [11.0] -5.9 [10.8] 0.6 [9.3]	**-25.7, 24.9** **-27.5, 15.7** **-20.6, 21.8**	**63.6** **57.7** **71.7**		Allison
Mexico 2015 *BT version NR*	815 236 343 236	All <10kg 10-18kg >18kg	-0.9 [15.9] 0.5 [20.8] -2.6 [12.1] -0.0 [15.1]	**32.1, 30.3** **-40.3, 41.3** **-26.3, 21.1** **-29.6, 29.6**	55.3 46.2 64.1 51.5		Khouli
New Zealand 2015 BT2011A	305	All			73.4		Britnell
India 2015 BT2007B	769	All	8.5 [14.6]	**-20.1, 37.1**	**43.8**		Asskaryar
India 2015	416	All	**0.3 [13.8]**	**-26.6, 27.3**	**53.1**		Asskaryar
South Africa 2016 BT2011A	300 73 128 99	All <12kg 12-20kg >20kg	5.5 [13.4] 7.3 [12.5] 6.3 [11.2] 3.3 [16.3]	**-20.8, 31.8** **-17.2, 31.8** **-15.7, 28.3** **-28.6, 35.2**	54.0 60.3 53.1 50.5	84.3 86.3 87.5 78.8	Georgoulas
Australia 2017 *BT version NR*	160 18 78 60 5	All <1yr 1-5yrs 6-10yrs 11-14yrs			68.2 50.0 73.1 54.4 91.4		O'Leary
South Africa 2017 BT2011A	328 30	All <12kg	-5.1 [16.2] 4.8 [8.9]	-36.9, 26.6 -12.7, 22.3	49.4 73.3	79.9 93.3	Wells

	90	12-20kg	-0.7 [12.3]	-24.8, 23.5	60.0	91.1	
	212	>20kg	-8.4 [17.4]	-42.5, 25.6	45.8	73.6	
	324	All	13.0		52.5	81.4	
USA 2015	65	<1yr	18.7				Chavez
BT version NR	163	1-5yrs	10.0				
	96	6-12yrs	14.8				
Nigeria 2016 *BT version NR*	300	All	3.1 [19.5]	**-35.2, 41.3**	**46.3**		Aliyu
Korea 2016 BT2011A	906	All	6.4 **[18.0]**	**-31.1, 39.5**	60.6		Jung
USA 2016 BT2011A	2434	All			54.4		Lowe
Japan 2016 BT2007B	237	All			75.9	97.5	Nosaka
Multinational 2016 BT2007B BT2011A	453990	All	4.6 [10.4]	**-15.8, 25.0**	62.9	95.6*	Ralston
Malaysia 2016 *BT version NR*	1163	All	2.6 [20.7]	**-37.9, 43.2**	**36.8**		Sahar
Zambia 2016 *BT version NR*	1282	All	4.3		47.3	75.1	Bowen
Thailand 2017	341	All	4.5 [15.4]	**-25.7, 34.7**	46.3	78.0	Samerchua

Devised Weight Estimating Method (DWEM)

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
USA 1986	258	All			61.0		Garland
		<20kg			66.0		
	484	All	**0.6 [11.5]**	**-22.0, 23.1**	**61.5**	**91.8**	
	121	<10kg	7.0 **[11.2]**	**-15.0, 29.0**	**54.1**	**86.9**	
Australia 2002	132	10-25kg	0.7 **[8.8]**	**-16.5, 17.9**	**74.3**	**97.7**	Black
	121	25-40kg	-2.0 **[12.3]**	**-26.1, 22.1**	**57.8**	**89.1**	
	110	>40kg	-3.9 **[10.7]**	**-24.9, 17.1**	**61.9**	**92.1**	
	400	All	**-2.8 [14.2]**	**-30.5, 25.0**	**51.0**	**83.3**	
	100	<10kg	0.4 [14.5]	**-28.0, 28.8**	**50.9**	**83.2**	
USA 2007	100	10-20kg	1.8 [10.6]	**-18.9, 22.5**	**64.8**	**93.7**	DuBois
	100	20-36kg	-6.7 [12.8]	**-31.8, 18.3**	**50.6**	**83.2**	
	100	>36kg	-6.6 [16.0]	**-37.9, 24.8**	**43.4**	**75.1**	

PAWPER Tape

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
	453	All	0.0 [6.2]	**-12.2, 12.2**	89.2	99.1	
	120	<12kg	0.9 [6.1]	**-11.1, 12.9**	90.8	100	
South Africa 2013	193	12-20kg	0.0 [5.9]	**-11.6, 11.6**	90.2	99.0	Wells
	140	>20kg	-0.9 [6.8]	**-14.2, 12.4**	86.4	98.7	
	300	All	1.6 [6.7]	**-11.5, 14.7**	87.7	98.3	
	73	<12kg	4.5 [7.6]	**-10.4, 19.4**	78.1	95.9	
South Africa 2016	128	12-20kg	1.0 [6.0]	**-10.8, 12.8**	93.0	99.2	Georgoulas
	99	>20kg	0.4 [6.2]	**-11.8, 12.6**	87.9	99.0	
	170	All			77.1		
	18	<1yr			55.6		
Australia 2015	78	1-5yrs			80.8		O'Leary
	64	6-10yrs			72.1		
	10	11-14yrs			88.6		
	328	All	1.1 [7.7]	-13.9, 16.1	83.4	98.5	
	30	<12kg	6.4 [6.5]	-6.2, 19.1	70.0	100	
South Africa 2017	90	12-20kg	1.2 [5.7]	-10.1, 12.4	90.0	100	Wells
	212	>20kg	0.3 [8.2]	-15.8, 26.5	82.5	97.6	
	324	All	10.1		63.0	89.2	
USA 2015	65	<1yr	13.1				Chavez
	163	1-5yrs	8.2				
	96	6-12yrs	11.5				
	1698	All	-2.6 **[12.1]**	**-26,3, 21.1**	63.5		
	542	<12kg	4.2 **[11.1]**	**-17.6, 26.0**	63.8		
USA 2015	611	12-20kg	-2.4 **[8.6]**	**-19.3, 14.5**	77.9		Garcia
	545	>20kg	-9.5 **[12.5]**	**-34.0, 15.0**	47.0		
USA 2017	13134	All	-0.9 [8.5]	**-17.5, 15.7]**	81.9	98.0	Wells

Mid-arm circumference (MAC) formula

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
	1370	All	2.9 **[17.5]**	-31.4, 37.2	44.2	85.1	
	449	1-5yrs	17.2 **[15.2]**	-12.5, 47.0	26.5	79.7	
Hong Kong 2010	473	6-8yrs	-0.5 **[14.1]**	-28.0, 27.1	53.9	88.6	Cattermole
	448	9-11yrs	-7.9 **[12.7]**	-32.9, 17.0	51.8	86.8	
	2130	All	15.7 **[21.1]**	-18.4, 49.8	33.5	56.7	
USA 2012	1169	1-5yrs	25.3 **[15.7]**	-5.1, 55.7	13.6	31.3	Cantle
	961	6-10yrs	4.1 **[24.6]**	-18.4, 26.6	57.8	87.6	
Botswana 2012	777	All			32.2		Wozniak
USA 2012	529	All	3.5 [12.1]	**-20.2, 27.2**	15.5	23.9	Abdel-Rahman
	324	All	28.8		24.1	40.1	
USA 2015	65	<1yr	>100				Chavez
	163	1-5yrs	33.2				
	96	6-12yrs	12.4				
Multinational 2016 BT2007B	453990	All	4.6 [10.4]	**-15.8, 25.0**	27.9	61.8*	Ralston

BT2011A

Mercy Method

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
USA 2012	1938	All	-0.5 [8.4]	**-16.8, 16.0**	78.6	98.0	Abdel-Rahman
USA 2013	624	All	1.7 [8.5]	**-15.0, 18.4**	76.4	98.1	Abdel-Rahman
USA 2013	976	All	-0.3 [9.4]	-18.7, 18.1	77.0	95.4	Abdel-Rahman
Mali 2014	473	All	1.6 [9.3]	**-16.6, 19.8**	71.5	96.7	Dicko
India 2014	374	All	1.5 [9.9]	**-17.9, 20.9**	70.0	96.0	Batmanabane
South Africa 2016	300	All	-0.9 [8.3]	**-17.2, 15.4**	80.1	98.2	Georgoulas
	73	<12kg	2.3 [10.6]	**-18.5, 23.1**	69.9	91.8	
	128	12-20kg	-1.5 [7.7]	**-16.6, 13.6**	82.0	100	
	99	>20kg	-2.6 [6.4]	**-15.1, 9.9**	84.8	100	
Australia 2015	199	All			69.7		O'Leary
	18	<1yr			66.7		
	78	1-5yrs			69.2		
	68	6-10yrs			72.1		
	35	11-14yrs			65.7		
Thailand 2017	430	All	-9.6 [10.9]		50.7	82.3	Samerchua
South Africa 2017	328	All	-6.7 [8.3]	-23.1, 9.6	63.9	94.3	Wells
	30	<12kg	-9.9 [11.3]	-32.1, 12.3	43.3	86.7	
	90	12-20kg	-7.3 [8.2]	-23.3, 8.7	67.8	91.1	
	212	>20kg	-6.0 [7.8]	-21.4, 9.4	65.1	96.7	

Wozniak method

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
Botswana 2012	777	MAC-ulna			80.8		Wozniak
		MAC-tibia			81.1		
South Africa 2016	300	All	-3.6 [15.7]	**-34.4, 27.2**	70.6	88.3	Georgoulas
	73	<12kg	-18.9 [22.4]	**-62.8, 25.0**	37.0	65.8	
	128	12-20kg	1.0 [9.8]	**-18.2, 20.2**	76.6	93.8	
	99	>20kg	1.0 [7.0]	**-12.7, 14.7**	87.9	98.0	
South Africa 2017	328	All	-4.0 [17.3]	-37.8, 29.9	64.5	84.0	Wells
	30	<12kg	-28.5 [17.1]	-62.0, 5.1	10.0	26.7	
	90	12-20kg	0.5 [23.2]	-45.1, 46.0	58.9	88.9	
	212	>20kg	-2.4 [10.2]	-22.3, 17.5	74.5	90.1	

Oakley System

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
Australia 2002	442	All	**5.4 [16.5]**	**-26.9, 37.7**	**43.4**	**75.0**	Black
	121	<10kg	16.0 **[14.0]**	**-11.4, 43.4**	**30.2**	**60.7**	
	132	10-25kg	10.0 **[9.4]**	**-8.4, 28.4**	**48.3**	**85.6**	
	121	25-40kg	-1.0 **[14.6]**	**-29.6, 27.6**	**50.6**	**82.8**	
	68	>40kg	-11.0 **[16.8]**	**-43.9, 21.9**	**37.1**	**67.1**	
New Zealand 2005	909	All	**-16.2 [16.9]**	**-49.2, 16.8**	**29.6**	**57.3**	Theron
	226	Maori	-15.8 [20.1]	**-55.1, 23.6**	**28.7**	**54.5**	
	420	Pacific	-19.8 [14.5]	**-48.1, 8.6**	**23.0**	**50.2**	
	160	European	-10.3 [15.1]	**-40.0, 19.3**	**40.3**	**71.7**	
	79	Asian	-10.8 [17.0]	**-44.0, 22.5**	**37.1**	**67.1**	

Growth Chart methods

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
England 2009	544	All	-4.0 **[12.5]**	-29.0, 20.0	**55.3**	**87.2**	Sandell
Thailand 2012	595	All			51.4		Trakulsrichai

Sandell tape

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
Australia 2014	2102	All	12.4 [11.7]	**-10.5, 35.3**	60.7		Allison
	363	<1yr	13.3 [13.7]	**-13.6, 40.2**	53.5		
	1739	1-5yrs	12.2 [11.0]	**-9.4, 33.8**	63.7		

Hanging-leg weight technique

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
USA 1990	100	All	9.0		73.7		Haftel
		<10kg			50.0		
		10-25kg			80.0		
		>25kg			90.0		

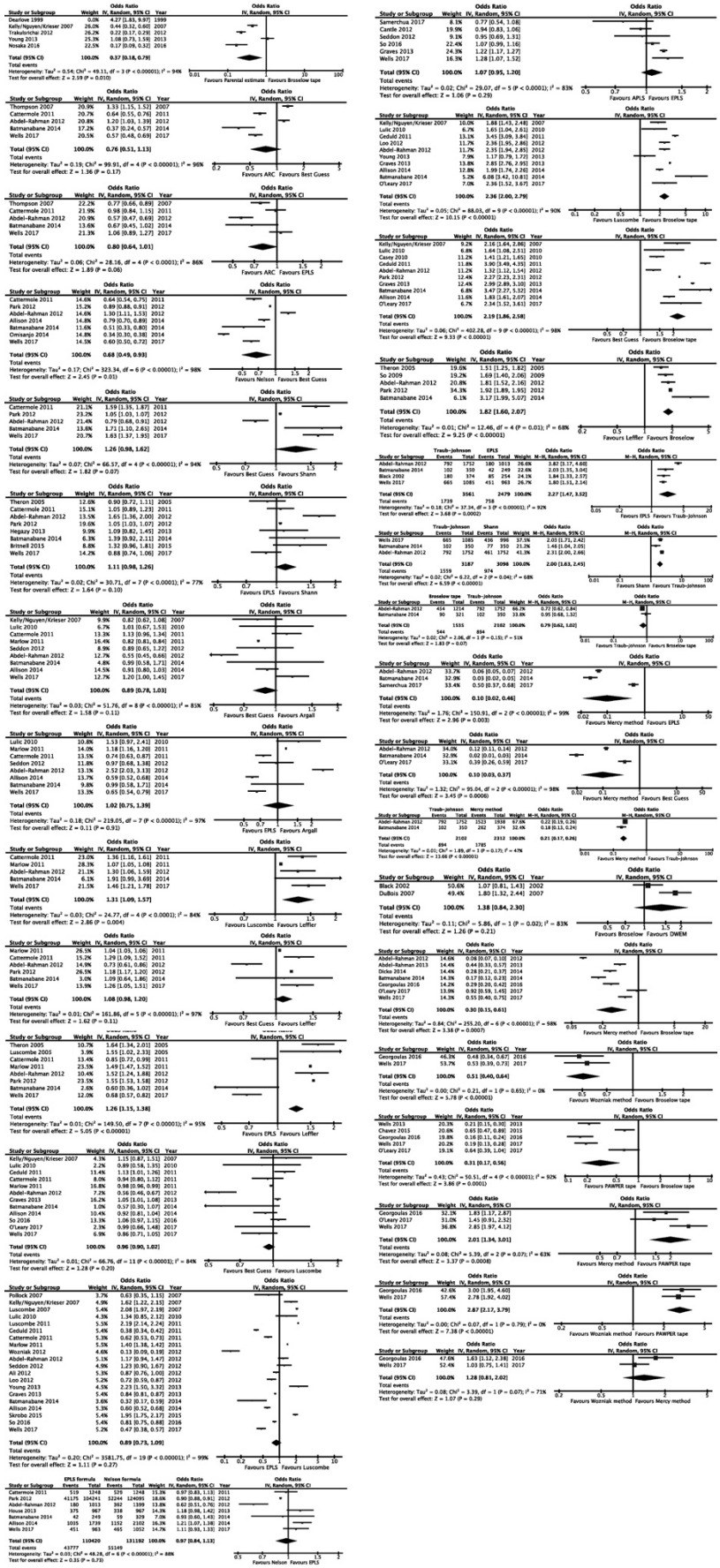
Finger Counting system

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
USA 2013	207	All			59.0	87.0	Young
Hong Kong 2016	4178	All	0.1 [14.8]	-34.0, 34.2	48.6	79.7	So

Traub-Johnson (TJ) formula

Country	N	Subgroup	MPE [SD] (%)	PELOA (%)	PW10 (%)	PW20 (%)	Author
Australia 2002	374	All	**-7.1 [13.6]**	**-33.6, 19.5**	**48.0**	**80.5**	Black
	132	10-25kg	-2.6 **[9.4]**	**-21.0, 15.8**	**69.4**	**96.0**	
	121	25-40kg	-4.0 **[13.5]**	**-30.5, 22.5**	**52.2**	**84.4**	
	121	>40kg	-15.0 **[14.0]**	**-42.4, 12.4**	**32.3**	**63.3**	
USA 2012	1752	All	-7.1 [15.8]	**-38.1, 23.9**	45.2	69.3	Abdel-Rahman
India 2014	350	All	11.1 [16.0]	**-20.3, 42.5**	29.0	62.0	Batmanabane
South Africa 2017	1085	All	-0.2 [13.2]	-26.2, 25.7	61.3	87.3	Wells

Traub-Kichen (TK) formula							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
Australia 2002	374	All	**-7.5 [14.4]**	**-35.8, 20.7**	**45.7**	**77.9**	Black
	132	10-25kg	-0.7 **[8.8]**	**-17.9, 16.5**	**74.3**	**97.7**	
	121	25-40kg	-6.5 **[12.3]**	**-30.6, 17.6**	**52.2**	**84.8**	
	121	>40kg	-16.0 **[16.8]**	**-48.9, 16.9**	**30.0**	**57.8**	
USA 2012	1747	All	-9.5 [15.9]	**-40.7, 21.7**	45.3	67.8	Abdel-Rahman
India 2014	344	All	10.0 [16.4]	**-22.1, 42.1**	28.0	63.0	Batmanabane
South Africa 2017	990	All	-0.8 [12.6]	-25.4, 23.9	62.5	88.8	Wells



Supplementary Figure 1 The full Review manager (Revman 5.3) figures for each direct comparison between weight estimation systems.

4.3 Epilogue

4.3.1 Major findings

This was the first study to propose a benchmark accuracy by which to evaluate the performance of weight estimation systems, based on a previous recommendation from a biostatistical research report [7]. This was used throughout the research in this thesis to set an acceptable outcome standard for weight estimation. It was founded on a tolerance level of <5% for critical errors (errors that would result in a drug dosing error of >20%). This benchmark may be considered too stringent by some authors, who argue that there is limited evidence to show harm from medication errors of 20% or even 30% [8]. Pragmatically, however, this proposed standard of accuracy has been achieved by the PAWPER tape and the Mercy method in many studies in different populations. It would be difficult to argue for a lower standard than that already achieved by existing systems unless some other compelling reason was advanced as justification.

The study found that age-based weight estimation was the least accurate of all the systems and showed the greatest variability between populations. The dual length- and habitus-based systems along with parental estimates were the most accurate of the methods evaluated. The limited available evidence allowing direct comparisons between the PAWPER tape and the Mercy method showed the PAWPER tape to be significantly more accurate. There was insufficient evidence to directly compare parental estimates with the PAWPER tape and Mercy method. The Broselow tape was significantly more accurate than age formulas but significantly less accurate than the PAWPER tape and the Mercy method.

4.3.2 This study in the context of the thesis

This study provided a useful starting point for this thesis to evaluate and objectively compare weight estimation systems. It provided key information that would be useful for clinicians and decision-makers, from both high- and middle- and low-income countries, about which weight estimation systems would be the most accurate to use. It also provided direct meta-analysis comparisons between different methods to stratify their performance.

This study verified that the PAWPER tape is one of the three most accurate weight estimation systems in the world, at present. It provided affirmation that the PAWPER methodology has merit and that further research to fine-tune the system would be appropriate.

The study did not address the appropriateness of any particular weight estimation system for use during medical emergencies – this will be dealt with in Chapter 10.

4.3.3 Link to the next article

In order to argue for the use of the modern dual length- and habitus-based systems in clinical practice, it was also necessary to specifically address the performance and current usage of the older weight estimation systems. The next article, found in Chapter 5, addresses age-based formulas specifically. Given the poor performance of age formulas in previous systematic reviews and in this meta-analysis, it was important to explicitly address the appropriateness of these systems as they are still widely used. The chapter addresses the broad question of whether it can be considered evidence-based practice to continue to use these formulas.

4.3.4 References

[1] Kelly A. Estimation of children's weight in medical emergencies. In: Preedy V, editor. Handbook of anthropometry: Physical measures of human form in health and disease. New York: Springer-Verlag; 2012. p. 1151-1161.

[2] Young KD, Korotzer NC. Weight estimation methods in children: a systematic review. Annals of Emergency Medicine. 2016;68(4):441-451 e10.

[3] Abdel-Rahman SM, Ridge AL. An improved pediatric weight estimation strategy. Open Medical Devices Journal. 2012;4:87-97.

[4] Wells M, Coovadia A, Kramer E, Goldstein L. The PAWPER tape: A new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. Resuscitation. 2013;84(2):227-232.

[5] Advanced Life Support Group. Advanced Paediatric Life Support: The Practical Approach (APLS) 6th Edition. New York: BMJ Books - Publisher: John Wiley & Sons (Wiley-Blackwell); 2016.

[6] Wells M, Goldstein L, Bentley A. A systematic review and meta-analysis of the accuracy of weight estimation systems used in paediatric emergency care in developing countries. African Journal of Emergency Medicine. 2017;7:S36-S54.

[7] Stewart D. Accuracy of the Broselow tape for estimating paediatric weight in two Australian Emergency Departments: University of Sydney; 2009 [Accessed 27 September 2017, Master of Biostatistics]. Available from: https://ses.library.usyd.edu.au//bitstream/2123/6265/1/Declan%20Stewart%20WPP%20final.pdf

[8] Cattermole GN, Graham CA, Rainer TH. Age-based weight prediction is imprecise. Emergency Medicine Journal. 2011;29(1):82-83.

Chapter 5

It is time to abandon age-based emergency weight estimation in children!

5.1 Prologue

5.1.1 Background

In the previous chapter, the relative accuracy of many different weight estimation systems was established. Along with healthcare provider guesses of weight, age-based formulas were the least accurate of all the weight estimation system options. This has been a consistent finding in studies which have compared age formulas against other methods of weight estimation [1-4]. Age-based weight estimation deserved special attention in this thesis because of the widespread use of these formulas in clinical practice and advanced life support courses.

Despite the lack of evidence supporting their accuracy, age-based formulas continue to be used and taught [5-7]. The reason for this is not clear, as no compelling arguments have been provided to justify their ongoing usage. The only advantages of the age formulas are that they require no equipment to be used and, theoretically, if prehospital emergency services personnel alert the Emergency Department staff with the age of an incoming paediatric patient, some advanced preparation of drugs and equipment is possible. However, these advantages are offset by the inaccuacy of the formulas as well as the well-documented difficulty of users recalling the formulas and applying the simple arithmetic correctly during emergencies [8-10]. Furthermore, other, better methods of providing forewarning by prehospital healthcare providers are available (such as the Broselow tape) [11].

One of the most recent concepts in age-based weight estimation, inspired by other two-dimensional weight estimation methods, was the development of the first age- and habitus-based formulas [12]. These formulas performed very well in the derivation study but had not yet been externally validated. It was important to evaluate this new proposed age- and habitus-based method, to establish whether it could be a viable option for weight estimation and whether it could compete with the

proven length- and habitus-based systems such as the PAWPER tape and Mercy method.

Many previous studies on age-based formulas have included an evaluation of two length-based weight estimation formulas: the Traub-Kichen and the Traub-Johnson formulas [13, 14]. Since these two length-based weight prediction formulas have often been analysed with the age-based formulas, they were evaluated in this study as well. They have previously been shown to be more accurate than the age formulas, including in the meta-analysis in Chapter 4, but have not been widely studied.

5.1.2 Purpose of the study

Since age formulas are so widely used, it was important to conclusively and specifically prove whether they have any evidence-based role in emergency weight estimation, at least in terms of accuracy.

There were several purposes for this study:

- to evaluate every previously published age-based formula to establish whether any formula could rival the degree of accuracy achieved with dual length- and habitus-based methods;
- to perform the same evaluation for every length-based weight estimation formula;
- to test the accuracy of age- and length-based formulas when combined with an assessment of habitus (i.e. age- and habitus-based formulas and length- and habitus-based formulas).

If this comprehensive evaluation of formulas failed to prove accuracy in every formula ever developed, it would provide additional support for the motivation to abandon these methods and adopt the newer weight estimation modalities, such as the PAWPER tape or the Mercy method. An additional benefit of this study would be to provide additional insight into the value of a visual assessment of habitus (used in a different system) in improving weight estimation accuracy.

The hypothesis was that the age-based methods would not achieve good results, since the available evidence does not support the biological validity of age as a good predictor of weight [15]. It was also anticipated that the length-based formulas would perform better than the age-based formulas but would not be as accurate as the PAWPER tape or Mercy method, primarily because the length formulas were based on growth data from the 1970s while the newer systems were developed from more recent data.

This study supplemented the meta-analysis in Chapter 4 in addressing Objective 1 (to describe the accuracy of commonly-used weight estimation systems in both high-income and low- and middle-income countries through the systematic review of previous studies and meta-analysis of data obtained from these studies).

5.2 Published paper

Wells M, Goldstein L, Bentley A. It is time to abandon age-based emergency weight estimation in children! A failed validation of 20 different age-based formulas. Resuscitation. 2017;116(7):73-83.

The PDF of the published article is included. The online-only supplementary material is appended after the PDF. The tables or figures have been retained in the formatting in which they are archived in the journal website.

Resuscitation 116 (2017) 73–83

Contents lists available at ScienceDirect

Resuscitation

journal homepage: www.elsevier.com/locate/resuscitation

Simulation and education

It is time to abandon age-based emergency weight estimation in children! A failed validation of 20 different age-based formulas☆

Mike Wells*, Lara Nicole Goldstein, Allison Bentley

Division of Emergency Medicine, Faculty of Health Sciences, University of the Witwatersrand, 7 York Road, Parktown, Johannesburg, South Africa

ARTICLE INFO

Article history:
Received 18 March 2017
Received in revised form 26 April 2017
Accepted 15 May 2017

ABSTRACT

Aim: When children's weight cannot be measured during medical emergencies, it must be estimated, as weight is required for many interventions. Age-based formulas are the oldest weight estimation systems in children, but have been shown to be very inaccurate in many studies. This aim of this study was to evaluate the accuracy of age- and length-based formulas and to see if a measure of body habitus could be used to improve the performance of these formulas.

Methods: This was an analysis of data from a sample of 1085 children aged from 1 month to 16 years, collected from four Emergency Departments in Johannesburg, South Africa. Basic demographic and anthropometric data was collected and each child had a visual assessment of body habitus, quantified as a habitus score. Weight estimates from 20 existing age-based formulas and two length-based formulas were then compared against measured weight to determine their accuracy. Age- and length-based, habitus-modified models were developed and similarly evaluated.

© 2017 Elsevier B.V. All rights reserved.

Introduction

The use of age-based weight estimation formulas in children is as outdated as the use of cricoid pressure during rapid sequence intubation. There are negligible mitigating factors for their use and no credible study has shown these formulas to achieve an acceptably accurate outcome.

Why is weight needed when managing paediatric emergencies?

Since drug dose calculation in children is almost always based on their body weight, one of the essential prerequisites for accurate drug dose calculation is an accurate measure of body weight [1,2]. When a child cannot be weighed, a precise, simple and rapid estimation of weight is needed, as the safety and effectiveness of therapeutic interventions ultimately depend on the accuracy of the weight estimation.

Age-based formulas

Age-based formulas are the oldest of the weight-estimation systems, with the Weech formula, the precursor of the Nelson formula, first appearing in a published report in 1956 [3]. These formulas have become amongst the most commonly researched and taught weight-estimation systems for use in emergencies [4]. There are more than 22 different age-based formulas that have been described to date (see Table 1). They are also amongst the most treacherous of the weight estimation systems because of their lack of accuracy and their susceptibility to faulty recall and calculation errors during emergencies [5–7]. A simple desk-top exercise with a weight-for-age growth chart can demonstrate that these formulas can underestimate weight by as much as 40% in children on the 97th centile, or overestimate weight by as much as 55% in children on the 3rd centile [8].

Are there alternatives to aged-based formulas?

The only benefits that formulas offer are that they require no equipment and may, in theory, allow advanced preparation if the age of a child being transported to the Emergency Department is known. Virtually every other weight estimation system studied has been shown to be more accurate than age-based formulas, except for healthcare provider guesses [13]. Since there are other better, highly accurate weight-estimation systems available, such as

☆ A Spanish translated version of the abstract of this article appears as Appendix in the final online version at http://10.1016/j.resuscitation.2017.05.018.

* Corresponding author.
E-mail addresses: mike.wells@emergencymedicine.co.za (M. Wells), drg666@gmail.com (L.N. Goldstein), bentley@wits.ac.za (A. Bentley).

http://dx.doi.org/10.1016/j.resuscitation.2017.05.018
0300-9572/© 2017 Elsevier B.V. All rights reserved.

Table 1
Weight estimation formulas evaluated in this study. The Traub-Johnson and Traub-Kichen formulas are *length*-based, but are often considered alongside the age-based formulas. The formulas are shown with a description of the applicability of each formula as well as the best and worst performances in previous research studies. The European Life Support (EPLS) formula is the old version of the APLS formula which was changed in 2011. The *a priori* model that was used to evaluate the effects of habitus-modification on existing formulas is also shown. The new CAWR-3 formulas and the recently described formulas for Japanese children are not included in the table. Z = age in years; z = age in months; X = length in centimetres.

Name	Formula	Restrictions and Limitations		Accuracy
Ali formula	$\boldsymbol{Wt} = (2.5 \times \boldsymbol{Z}) + 8$	Derived in a Trinidadian population of children ≤5yrs of age in 2012. No validation studies to date. Age restriction 1–5 years of age.		No previous validation studies. PW10 47.5% in originator study [9].
Argall formula	$\boldsymbol{Wt} = (3 \times \boldsymbol{Z}) + 6$ **or** $[\boldsymbol{Wt} = 3 \times (\boldsymbol{Z} + 2)]$	Developed from a small UK study in 2003 (300 children). Mostly found to underestimate weight, more so in older and heavier children. Age restriction 1–10 years of age.		Best PW10: 46.5% in Australia (Allison 2014) [10] Worst PW10: 10.0% in India (Batmanabane 2014) [11]
Advanced Paediatric Life Support formula (new)	$\boldsymbol{Wt} = \frac{z}{2} + 4$ $\boldsymbol{Wt} = (2 \times \boldsymbol{Z}) + 8$ **or** $[\boldsymbol{Wt} = 2 \times (\boldsymbol{Z} + 4)]$ $\boldsymbol{Wt} = (3 \times \boldsymbol{Z}) + 7$	For infants ≤12 months of age. For children aged 1–5 years. For children aged 6–12 years.	Derived in a UK population and adopted in 2011 by the Advanced Life Support Group from a combination of the original APLS and the Luscombe formulas. It was untested and not yet validated at the time of adoption. In general overestimates weight. Age restriction birth to 12 years of age.	Best PW10: 36.4% in USA (Chavez 2015) [12] Worst PW10: 15.0% in Mali (Dicko 2014) [13]
Australian Resuscitation Council formula	$\boldsymbol{Wt} = 3.5$ $\boldsymbol{Wt} = (2 \times \boldsymbol{Z}) + 8$ $\boldsymbol{Wt} = 3.3 \times \boldsymbol{Z}$	At birth. For children aged 1–9 years. For children 10 years and over.	Adopted by the ARC in Australia in 1996. Mostly underestimates weight, more so in older and heavier children. Differing accuracy in different ethnic, socio-economic and international populations. No specific age restriction noted.	Best PW10: 42.1% in Hong Kong (Cattermole 2011) [14] Worst PW10: 23.0% in India (Batmanabane 2014) [11]
Best Guess formulas	$\boldsymbol{Wt} = \frac{z+9}{2}$ $\boldsymbol{Wt} = (2 \times \boldsymbol{Z}) + 10$ **or** $[\boldsymbol{Wt} = 2 \times (\boldsymbol{Z} + 5)]$ $\boldsymbol{Wt} = 4 \times \boldsymbol{Z}$	For infants ≤12 months of age. For children aged 1–5 years. For children aged 6–14 years.	Also known as the Tinning formulas. Derived in Australian population in 2007 from a retrospective database study of more than 70 000 children. Frequently overestimates weight, especially in poorer populations. Has been evaluated in several follow-up studies with mixed results.	Best PW10: 48.9% in Australia (Allison 2014) [10] Worst PW10: 10.0% in India (Batmanabane 2014) [11]
Bicer formula	$\boldsymbol{Wt} = (3 \times \boldsymbol{Z}) + 6$ **or** $[\boldsymbol{Wt} = 3 \times (\boldsymbol{Z} + 2)]$ $\boldsymbol{Wt} = (4 \times \boldsymbol{Z}) - 3$	For children aged 3–6 years. For children aged 7–18 years.	Although these formulas are mentioned and evaluated in the Bicer study [15], the analysis and reporting is very poor; the origin of the first formula of the set is the same as the Argall formula. Age restriction proposed by Bicer to be 3–18 years.	No previous validation studies.
Chinese age-weight rule 1	$\boldsymbol{Wt} = (3 \times \boldsymbol{Z}) + 5$	For children aged 1–10 years.	Developed in Hong Kong for ethnic Chinese children in 2011 from a sample of 1248 children. Age restriction 1–10 years (although developers advise use with caution over 7 years).	Best PW10: 36.1% in Hong Kong (Cattermole 2011) [14] Worst PW10: 30.9% in USA (Cantle 2012) [16]
Chinese age-weight rule 2	$\boldsymbol{Wt} = \frac{(\boldsymbol{Z} \times 7) + 25}{3}$ $\boldsymbol{Wt} = (4 \times \boldsymbol{Z}) - 4$	For children aged 1–6 years. For children aged 7–10 years.		Best PW10: 44.7% in USA (Cantle 2012) [16] Worst PW10: 38.7% in Hong Kong (Cattermole 2011) [14]

Erker formulas	$\boldsymbol{Wt} = (2 \times \boldsymbol{Z}) + 6$ $\boldsymbol{Wt} = (3 \times \boldsymbol{Z}) + 6$ $\boldsymbol{Wt} = (4 \times \boldsymbol{Z}) + 6$	For "tall n' thin" children For "normal" children For "tiny n' thick" children	Developed in Germany, derived from weight-for-age percentiles [17]. Makes use of a visual assessment of habitus to select which formula must be used. It is the only habitus-modified age-based formula method. Has not yet been validated in a study on actual children. For children 1–12 years.	No previous validation studies. In this study, two models were applied to the Erker formulas. In Erker 1 the "tall n' thin" formula was used in children with HS1 and HS2, the "normal" formula was used in children with HS3 and the "tiny n' thick" formula used for children with HS4 and HS5. In Erker 2 the differences were that the "normal" formula was used in children with HS3 and HS4 and the "tiny n' thick" formula was used for children with HS5. No other variations resulted in improved outcomes.
European Paediatric Life Support formula	$\boldsymbol{Wt} = 2 \times (\boldsymbol{Z} + 4)$ ***or*** $[\boldsymbol{Wt} = (2 \times \boldsymbol{Z}) + 8]$	Original population and date of derivation unclear. Often underestimates weight, more so in older and heavier children. Differing accuracy in different ethnic, socio-economic and international populations. Age restriction 1–10 years of age.		Best PW10: 59.5% in Australia (Allison 2014) [10] Worst PW10: 22.2% in Australia (Luscombe 2005) [18]
Finger counting	$\boldsymbol{Wt} = (2.5 \times \boldsymbol{Z}) + 7.5$	Although not actually used as a formula, the finger counting method approximates this age-based formula. The account of its origin differs between different articles.		Best PW10 59.0% in originator study [1].
Garwood formula	$\boldsymbol{Wt} = \frac{z}{4} + 6$	Developed in a UK population from a sample of 1252 children in 2012. The initial validation study was fatally flawed, but this formula has been subjected to a follow-up study subsequently (showing poor performance). For children aged 1 to 16yrs.		Only PW10: 35.8% in Egypt (Hegazy 2013) [2]
Leffler formulas	$\boldsymbol{Wt} = \frac{Z+8}{2}$ $\boldsymbol{Wt} = (2 \times \boldsymbol{Z}) + 10$	For children <1 year of age. For children aged 1–10 years.	Also known as the Tintinalli formulas, the original origin is unclear, but became popular after the Leffler study in 1997. Overestimates weight in younger children (≤6yrs) and underestimates weight in older children (> 6yrs).	Best PW10: 46.2% in USA (Leffler 1997) [3] Worst PW10:11.0% in India (Batmanabane 2014) [4]
Luscombe formula	$\boldsymbol{Wt} = (3 \times \boldsymbol{Z}) + 7$	Developed in the UK in 2007 from a large database of nearly 14 000 children. Underestimates weight in most populations studied, but significantly overestimates weight in populations from developing countries. Age restriction 1–10 years.		Best PW10: 46.8% in Australia (Allison 2014) [5] Worst PW10: 7.2% in Botswana (Wozniak 2012) [6]
Michigan formula	$\boldsymbol{Wt} = (3 \times \boldsymbol{Z}) + 10$	Derived by Ackwerh in 2010 [7], but has only been evaluated in Zambia.		Best PW10: – Worst PW10: 4.0% in Zambia (Bowen 2016) [8]
Nelson formulas (originally Weech's formulas)	$\boldsymbol{Wt} = \frac{Z+9}{2}$ $\boldsymbol{Wt} = 2 \times (\boldsymbol{Z} + 4)$ $\boldsymbol{Wt} = \frac{(z \times 7) - 5}{2}$	For infants 3–12 months. For children aged 1–6 years. For children aged 7–12 years.	As described in Nelson's Textbook of Paediatrics. The origin is probably from Weech's formulas, first reported in 1954 in the USA. The Weech formula is still in use today as one of the standard measurement denominators for determining underweight status. Weight most often overestimated in infants and older children (> 6yrs) and underestimated in younger children (≤6 years).	Best PW10: 54.8% in Australia (Allison 2014) Worst PW10: 18.0% in India (Batmanabane 2014) [4]

Table 1 (*Continued*)

Name	Formula	Restrictions and Limitations		Accuracy
Park formulas	$\boldsymbol{Wt} = \frac{z+9}{2}$ $\boldsymbol{Wt} = (2 \times \boldsymbol{Z}) + 9$ $\boldsymbol{Wt} = (4 \times \boldsymbol{Z}) - 1$	For infants ≤12 months of age. For children aged 1–4 years. For children aged 5–14 years.	Derived in a Korean population from a large database study (nearly 125 000 children). Poor accuracy in older children (> 6 years).	Best PW10: 41.7% in originator study [9] Worst PW10: 21.9% in Zambia (Bowen 2016) [8]
Shann formulas	$\boldsymbol{Wt} = (2 \times \boldsymbol{Z}) + 9$ $\boldsymbol{Wt} = (3 \times \boldsymbol{Z})$	For children aged 1–9 years. For children aged >9 years.	Used in Australasia primarily. Its origin is unclear. Underestimates weight increasingly with increasing age.	Best PW10: 53.1% in New Zealand (Britnell 2015) [10] Worst PW10: 22.0% in India (Batmanabane 2014) [4]
Theron formula	$\boldsymbol{Wt} = \boldsymbol{e}^{(0.175571 \times \boldsymbol{Z}) + 2.197099}$	Derived in 2005 in New Zealand from a small study of 900 children that included many Pacific Island children (high weight-for-age). The developers intended it for use in children high in the weight-for-age centiles. Age restriction 1–10 years. Overestimates weight in most populations.		Best PW10: 27.5% in New Zealand (Britnell 2015) [10] Worst PW10: 7.0% in India (Batmanabane 2014) [4]
Unnamed formula	$\boldsymbol{Wt} = (3 \times \boldsymbol{Z}) + 8$	This formula is mentioned in some weight-estimation studies with no reference to its origin. It is not known what restrictions apply. Mentioned by Dearlove [11].		No previous validation studies.
Traub-Johnson formula	$\boldsymbol{Wt} = 2.05 \times \boldsymbol{e}^{0.02\boldsymbol{X}}$	Derived in 1980 from USA national growth data from 1959, this formula was used to estimate ideal body weight. The formula was intended to estimate the 50th centile of weight-for-height. For children aged 1–18 years.		Best PW10: 48.0% in Australia (Black 2002) [12] Worst PW10: 29.0% in India (Batmanabane 2014) [4]
Traub-Kichen formula	$\boldsymbol{Wt} = 2.396 \times 1.0188^{\boldsymbol{X}}$	Derived in 1983 in the USA from data from more than 20000 children in the National Centre for Health Statistics database. The formula was intended to estimate ideal body weight. For children over 74 cm and aged 1–17 years.		Best PW10: 45.7% in Australia (Black 2002) [12] Worst PW10: 28.0% in India (Batmanabane 2014) [4]
A priori models	**[existing EPLS formula]**	For HS1 decrease weight estimate by 20%, for HS2 decrease weight estimate by 10%, for HS3 leave weight estimate unchanged, for HS4 increase weight estimate by 10%, for HS5 and above increase weight estimate by 25%.		*A priori* models for evaluating the effects of habitus-modification of existing formulas on the accuracy of the prediction.
Derived models	$\boldsymbol{HS}1 : \boldsymbol{Wt} = (1.9 \times \boldsymbol{Z}) + 5.8$ $\boldsymbol{HS}2 : \boldsymbol{Wt} = (2.3 \times \boldsymbol{Z}) + 5.8$ $\boldsymbol{HS}3 : \boldsymbol{Wt} = (2.4 \times \boldsymbol{Z}) + 7.5$ $\boldsymbol{HS}4 : \boldsymbol{Wt} = (2.9 \times \boldsymbol{Z}) + 8.3$ $\boldsymbol{HS}5 : \boldsymbol{Wt} = (3.7 \times \boldsymbol{Z}) + 8.1$	The derivative model with the best outcome.		

Table 2
Characteristics of the study population. Cut-off limits for definition of overweight were a BMI-for-age centile of >85% (Z-score >+1.44), for obesity a BMI-for-age centile of >95% (Z-score >+2) and for underweight a BMI-for-age centile of <5% (Z-score <-2). The age-group categories were based on those most often used in weight-estimation studies. Of the total study population, 31.8% fell below the 25th centile of BMI-for-age, 29.6% above the 75th centile and 38.6% within the interquartile range.

	All	< 1yr	1 to 5yrs	5 to 10yrs	> 10yrs
Number (%)	1085	89 (8.2)	457 (42.1)	413 (38.1)	126 (11.6)
Male (%)	561 (51.7)	50 (56.2)	248 (54.3)	195 (47.2)	68 (54.0)
Age (yrs) median (IQR)	4.9 (2.3, 7.7)	7 m (6 m,10m)	2.9 (1.9, 4.0)	7.1 (5.9, 8.4)	11.3 (10.6, 12.2)
Weight (kg) median (IQR)	18.2 (13.0, 25.7)	8.1 (6.9, 9.0)	14.0 (11.7, 16.2)	23.3 (20.0, 28.6)	39.0 (30.7, 45.6)
BMI z-score median (IQR)	0.0 (−1.0, 0.9)	0 (−0.7, 0.8)	−0.2 (-1.1, 0.8)	0.1 (−0.8, 0.9)	0.1 (−0.8, 0.9)
Overweight (%)	126 (11.6)	7 (8.2)	49 (10.7)	36 (8.7)	15 (11.9)
Obese (%)	62 (5.7)	4 (4.1)	27 (6.0)	24 (5.8)	10 (7.9)
Overweight or obese (%)	188 (17.3)	11 (12.2)	76 (16.6)	60 (14.5)	25 (19.8)
Underweight (%)	91 (8.4)	8 (9.2)	27 (6.0)	49 (11.8)	13 (10.3)
HS1	63 (5.8)	2 (2.2)	19 (4.2)	28 (6.8)	14 (11.1)
HS2	241 (22.2)	10 (11.2)	100 (21.9)	107 (25.9)	24 (19.0)
HS3	460 (42.3)	52 (58.4)	214 (46.8)	159 (38.5)	35 (27.8)
HS4	245 (22.6)	23 (25.8)	105 (23.0)	90 (21.8)	27 (21.4)
HS ≥ 5	76 (7.1)	2 (2.2)	19 (4.2)	29 (7.0)	26 (20.6)

parental estimates of weight, the Mercy method and the PAWPER tape, it is difficult to justify the continued use of age formulas.

Is there a way of improving age-based formulas?

Given that the poor performance of these age formulas is primarily because of the wide variability of weight-for-age, the inclusion of a modifier based on body habitus could, in theory, improve the accuracy of the prediction. There are several techniques that have proven the concept of length-based, habitus modified weight estimation: the Derived Weight Estimating Method (DWEM), the Yamamoto Obesity Icon system, the PAWPER tape, the Mercy method and, most recently, the Broselow tape weight modified according to waist circumference [14–18]. Although the statistical relationship between age and weight is less linear and less close than that between length and weight, age-based weight estimation could potentially be improved by taking habitus into account. This has already been shown in a theoretical study, but no validation studies have proven the concept [19].

Our primary objective was to evaluate the accuracy of every published weight-estimation formula in our study population (both age- and length-based). Our secondary objective was to evaluate new age- or length-based, habitus-modified models to see if they could achieve acceptable accuracy.

It was our hypothesis that no age-based formula would perform acceptably well, including the newest age-based, habitus-modified formulas and that, furthermore, we would not be able to derive an age-based, habitus modified formula that could perform as well as other contemporary habitus-modified systems.

Methods

This was a secondary analysis of pooled, prospectively collected data from three previous studies evaluating the PAWPER tape, under Human Research Ethics Committee of the University of the Witwatersrand approval M1511107 [16,20,21].

These studies were prospective, cross-sectional studies from four Emergency Departments (EDs) in Johannesburg, South Africa. Two of the EDs serve a community of mostly low socio-economic status and two serve a middle-class community. The first study recruited 453 children between September and October 2008 in two centres; the second study recruited 332 children between July and December 2014 in one centre; the third study recruited 300 children in a single centre between August 2014 and January 2015.

These studies made use of a convenience sample of children aged one month to 16 years who were not in need of urgent medical care. The exclusion criteria were children with congenitally abnormal stature and children whose length could not be accurately measured.

Methods and measurements

Study procedure: data was collected by doctors or professional nurses in the ED who had received brief instruction in the assessment of body habitus. A similar procedure was followed in each study:

- Basic demographic data, including age, was collected.
- Each child's supine length was measured with a standard measuring tape.

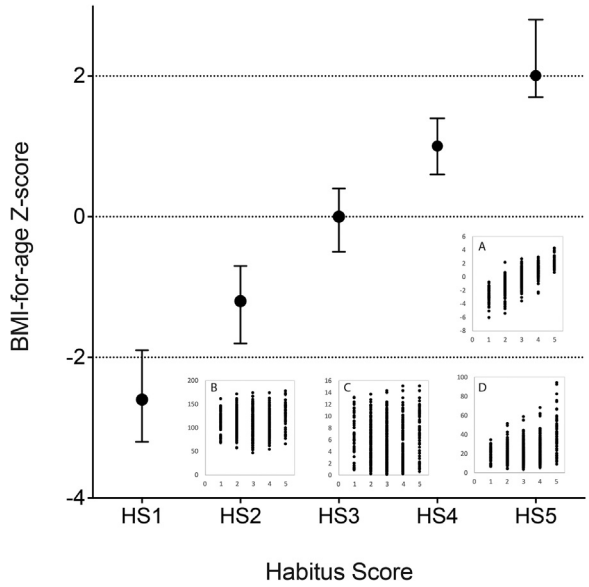
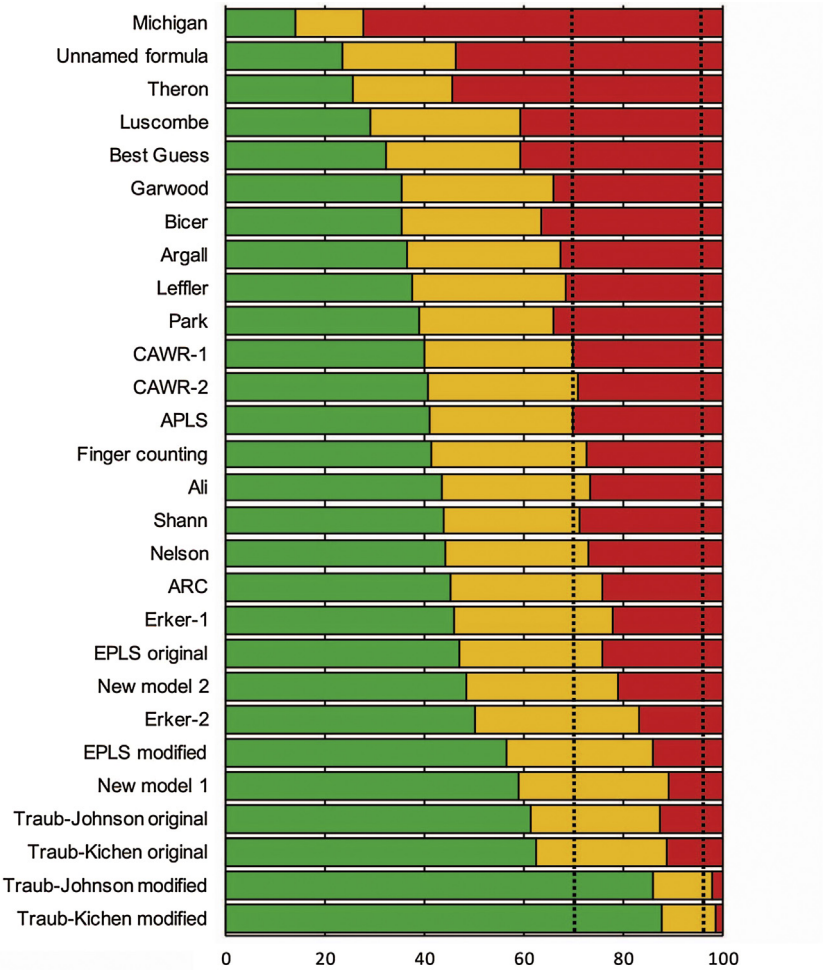


Fig. 1. The relationship between habitus score and anthropometric variables. The main chart illustrates the relationship between habitus score and BMI-for-age Z-score, with medians and interquartile ranges indicated. The median for each score was significantly different to its neighbours (Mann-Whitney, $p < 0.0001$). Insert A shows a scatterplot, illustrating the distribution of outliers for each habitus score category. Inserts B and C show scatterplots of length and age against habitus score. These illustrate the lack of discernible relationship between habitus score and these variables. Insert D shows a scatterplot of weight against, with limited association between the two other than the children with the heaviest weights were in the higher habitus score categories.



	N	MPE	LLOA	ULOA	RMSPE	PW10	PW20
Michigan formula	963	32.3	-19.1	83.7	34.3	13.9	27.5
Unknown formula	963	21.0	-26.4	68.3	24.7	23.5	46.1
Theron formula	1085	29.2	-43.9	102.4	32.8	25.4	45.4
Luscombe formula	883	15.5	-30.0	60.9	20.5	29.1	59.1
Best Guess formula	1075	16.1	-32.0	64.2	21.1	32.3	59.2
Garwood formula	996	9.2	-36.5	54.8	18.0	35.3	65.9
Bicer formula	751	11.4	-35.6	58.3	18.9	35.4	63.5
Argall formula	883	9.5	-34.9	53.9	17.4	36.5	67.3
Leffler formula	972	7.7	-35.3	50.7	16.9	37.6	68.3
Park formula	1075	10.9	-36.0	57.8	18.3	39.0	66.0
CAWR-1 formula	963	3.9	-40.4	48.2	16.4	39.9	69.6
CAWR-2 formula	883	7.9	-33.3	49.1	15.9	40.7	70.8
APLS formula	1066	6.1	-40.4	52.7	16.9	40.8	69.8
Finger counting formula	963	5.5	-35.5	46.5	15.5	41.4	72.5
Ali formula	478	9.2	-23.8	42.3	14.9	43.3	73.2
Shann formula	996	1.5	-39.7	42.7	15.2	43.8	71.3
Nelson formula	1052	4.0	-38.6	46.6	15.4	44.2	73.0
ARC formula	996	-2.0	-40.9	36.9	14.4	45.3	75.6
Erker formula 1	963	6.8	-28.6	42.2	14.0	45.8	77.9
EPLS formula	963	-4.2	-41.9	33.5	14.2	46.8	75.8
New model 2	635*	7.5	-24.1	39.1	13.1	48.2	78.7
Erker formula 2	963	2.1	-30.8	35.0	12.3	50.1	83.0
EPLS habitus modified	963	-5.2	-32.9	22.5	11.0	56.5	85.7
New model 1	635*	2.5	-25.8	30.7	10.3	58.7	89.1
Traub-Johnson formula	1085	-0.2	-26.2	25.7	10.0	61.3	87.3
Traub-Kichen formula	990	-0.8	-25.4	23.9	9.6	62.5	88.8
Pre-determined threshold for minimum acceptable accuracy of weight estimation – PW10 of ≥70% and PW20 of ≥95%							
Traub-Johnson habitus modified	1085	-0.7	-16.7	15.2	5.9	85.7	97.9
Traub-Kichen habitus modified	990	-1.3	-14.3	11.7	5.1	87.5	98.6

Fig. 2. Outcomes of the analyses of performance of the existing formulas and new habitus-modified, aged-based models. The green bar indicates the PW10, the red bar the critical errors (error >20%) and the orange bar the intermediate errors. The sample sizes included for each formula are different because of the different age limitations for each formula. A positive MPE indicates an overestimation of weight and a negative MPE an underestimation of weight. The wide PELOA shows a very poor precision of the age-based methods. The length-based habitus-modified formulas achieved an error of less than 15% in 95% of the population. No unmodified formula performed with adequate accuracy. The Erker formula performed slightly better than any of the existing unmodified formulas (paired *t*-test, McNemar test, $p < 0.0001$ for each test). The unmodified length-based formulas (Traub-Johnson and Traub-Kichen formulas) were significantly more accurate, with a smaller bias and higher precision than any of the age-based formulas, with the sole exception of one of the new habitus-modified models (paired *t*-test, McNemar test, $p < 0.0001$ for each test). The habitus-modified

Each child was assessed using a visual inspection to assign a numerical quantification of body habitus (a habitus score (HS) from 1 to 5 with HS1 representing a very thin or emaciated child, HS2 a slim or petite child, HS3 a child with optimum habitus, HS4 an overweight child and HS5 an obese child). The doctors and professional nurses who collected data for this study underwent a 20-min training session on how to assess body habitus, conducted by one of the authors (MW). This training focused on a three-step process: firstly, to visually inspect the child's body frame size and degree of adiposity and obtain a gestalt impression of their habitus. The raters were taught to assess each child as being of average habitus (HS3) or greater than average weight-for-length (HS4 or HS5) or less than average weight-for-length (HS1 or HS2) based on their gestalt impression. Finally, they would need to decide whether any deviation from average was moderate (HS2 or HS4) or more extreme (HS1 or HS5) to assign the appropriate habitus score value. Raters assigned habitus score values based on their gestalt perception rather than any anthropometric measurements or specific morphological features. Reference images were also developed and validated to increase the objectivity of habitus assessment (see supplementary figure) [16].

- The child's actual weight was then measured on a scale and recorded to the nearest 0.1 kg.

Inter-observer reliability: 134 children had habitus scores assessed by two researchers to determine inter-rater reliability.

Data extraction data from each of the contributing studies were pooled for the *post hoc* analysis. The data that were extracted from each of the 1085 children in this analysis included: age, BMI, habitus score and actual measured weight.

Model development: *a priori* models were devised before the analysis. Since the EPLS formula most closely predicted the 50th centile of weight-for-age and the two length-based formulas closely predicted the 50th centile of weight-for-length, they were chosen as the basis for the habitus modified models. The weight prediction models are shown at the bottom of Table 1. The single existing age- and habitus-based formula (the Erker formula, published in 2014) was calculated in two ways for further analysis (also shown in Table 1).

To establish the maximum potential effect of habitus modification of age formulas, additional new models were also explored using linear regression techniques on a derivation sample of 450 patients, with age, habitus score and sex evaluated as independent variables. The derived formulas were then subsequently tested in a separate validation sample of 635 children.

Analysis

Since actual weight, and estimated weight, varied greatly over the range of children in the study, a percentage error analysis was preferred over actual differences. Three statistical measures were used to assess the performance of the formulas:

- Mean percentage error (MPE) was calculated for each model as a quantification of the overall estimation bias.
- The 95% limits of agreement of the mean percentage error (PELOA) were calculated to represent a quantification of the estimation precision.
- The percentage of weight estimations that fell within 10% (PW10) and 20% (PW20) of actual weight were calculated for each system to quantify overall accuracy. An error of greater than 20% regarded as a critical error.

Subgroup analyses were performed for children in four age categories (<1 year, 1–5 years, 5–10 years and >10 years) and three habitus categories based on BMI-for-age Z-scores ("thin" children Z-score <-0.67, "fat" children Z-score >0.67 and "average" children in between).

The data was analysed to evaluate differences between the formula-predicted weight and actual weight using a comparison of MPE and PW10 analysis with parametric (paired t) and categorical (McNemar) tests respectively. A $p < 0.05$ was considered to be significant for all statistical tests.

Software: statistical analysis was performed using Stata (StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP.

The primary outcome measure was the performance of each of the age-based weight-estimation models when compared to actual measured weight. A PW10 >70% and a PW20 >95% (the level of accuracy achieved by the newest dual length-based, habitus-modified weight-estimation systems and one which has previously been suggested as a benchmark) was defined as an acceptable outcome for any formula [22].

Results

The basic demographic information is shown in Table 2. The study population was well distributed across the age- and weight-spectrum of children for whom the formulas could be used and had a diverse racial and ethnic makeup.

Accuracy of habitus assessment

The distribution of BMI-for-age Z-scores and the relationship between Z-scores and assigned habitus score categories for the study population are shown in Fig. 1. The inter-rater agreement for the assessment of body habitus, evaluated in a subsample of 134 children, showed a percentage agreement of 82.1%, a Cohen's kappa of 0.74 (0.64, 0.83) and a linear-weighted kappa of 0.81 (0.74, 0.89), which indicated a good to very good strength of agreement.

Linear regression models identified age and habitus score as the significant independent variables. The derived formulas are shown in Table 1.

Evaluation of existing and new formulas

The results of the analyses and sub-group analyses are shown in Fig. 2, Fig. 3a and b. The performance of weight estimation by the age-formulas was very poor, but was slightly better with length-based formulas. There was a very high incidence of critical errors (>20% error) with all age formulas. There was a consistent pattern across age categories with the best performance of formulas in the 1–5 years age group, decreasing with increasing age. No age formula performed well at any age, however. Variation in performance by body habitus category was better or worse in fat or thin children depending on when and where the formula was derived (e.g. the EPLS was better in underweight children than in overweight children while the opposite was true for the Luscombe formula). No age formula performed well in any habitus category, even in "average" habitus children. The unmodified length-based formulas showed a

length-based formulas significantly better than any of the other systems, including the habitus-modified age-based systems and surpassed the predetermined accuracy outcome criteria (paired *t*-test, McNemar test, $p < 0.0001$ for each test). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

similar pattern, but with exceptionally, near-perfect performance in the group of "average" children.

The effects of habitus-modification on performance

The Erker formulas, the only pre-existing habitus-modified formulas, performed better and were more consistent across the subgroups than the other age formulas, but did not achieve an acceptable level of accuracy. The *a priori* models and the derived models also failed to achieve a meaningful improvement in weight-estimation. The habitus-modified formulas still showed a decreasing accuracy with age, but were much more consistent across the body-types than the unmodified formulas. On the other hand, habitus-modification improved the performance of the Traub-Johnson and Traub-Kichen length-based formulas prodigiously, with excellent performance at all ages and in all body

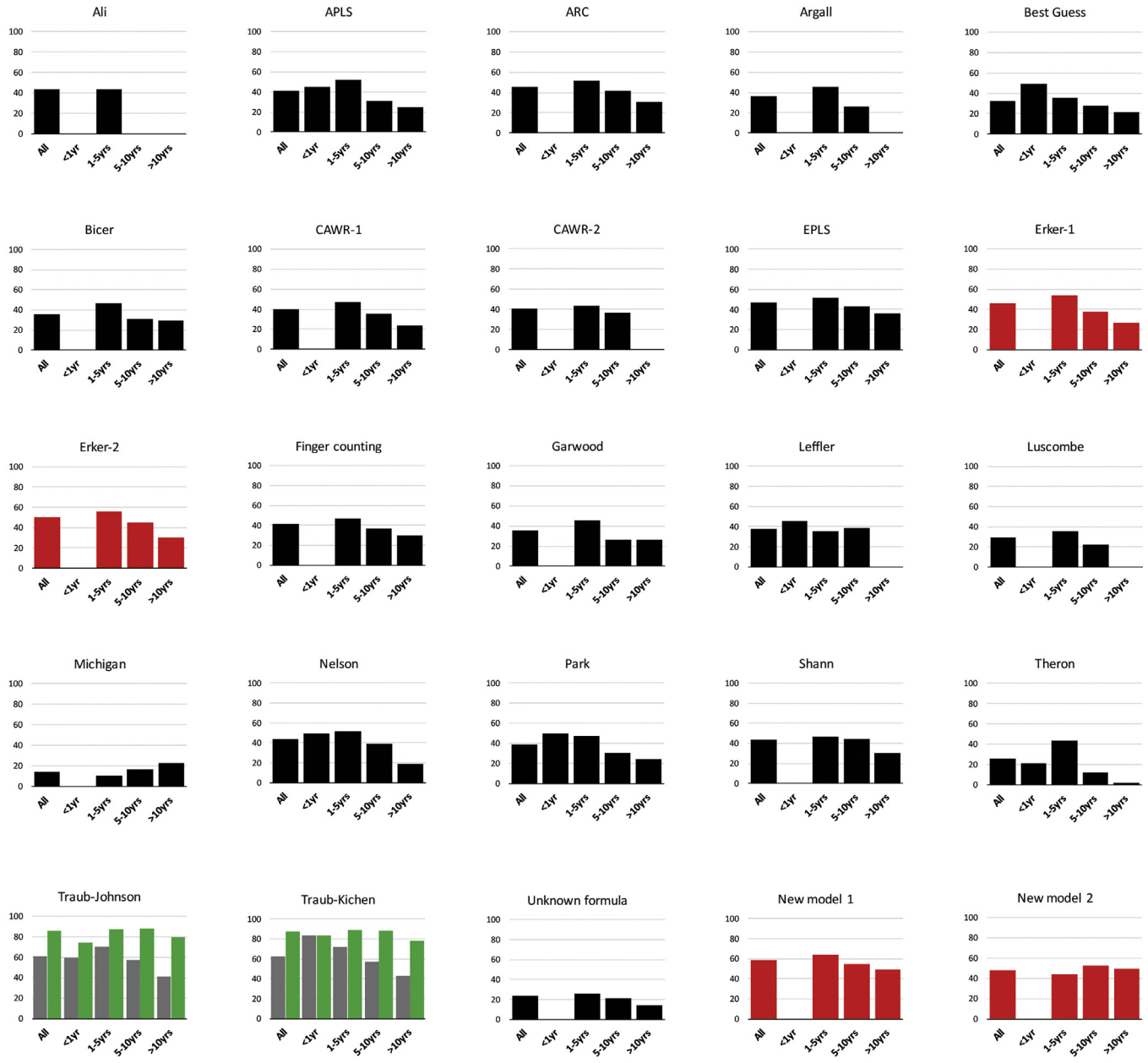


Fig. 3. (a) Aaccuracy data (PW10) for each formula with age subgroup outcomes. The pattern of highest accuracy in the 1–5-year age group, with decreasing accuracy in older children, is virtually uniform across the formulas, including the Erker habitus-modified formulas. The only exceptions were the newly derived models which showed a more consistent level of accuracy across the age-groups. The data for the length-based formulas is shown for both the unmodified and the habitus-modified formulas. The unmodified length-based formulas performed significantly better than the age-based formulas in every age category and the habitus-modified length-based formulas far outperformed all other systems (McNemar test, $p < 0.0001$). (b) Abccuracy data (PW10) for each formula with habitus subgroup outcomes. "Fat" children were defined as a BMI-for-age Z-score of >0.6745 (75th centile), "Thin" children were defined as a BMI-for-age Z-score of <0.6745 (25th centile) and "Average" children in the range between. There was widely differing accuracy between habitus subgroups depending on whether the formula was biased towards a low or high weight-for-age. This difference was far less apparent in the existing habitus modified system (Erker formula) and the new habitus modified models. The data for the length-based formulas is shown for both the unmodified and the habitus-modified formulas. The length-based formulas performed significantly better than the unmodified age-based formulas overall and in the "Average" and "Thin" categories (McNemar test, $p < 0.0001$). These formulas achieved a near-perfect accuracy performance in children in the "Average" category. The habitus-modified length-based formulas outperformed the other systems and performed consistently well across the habitus categories.

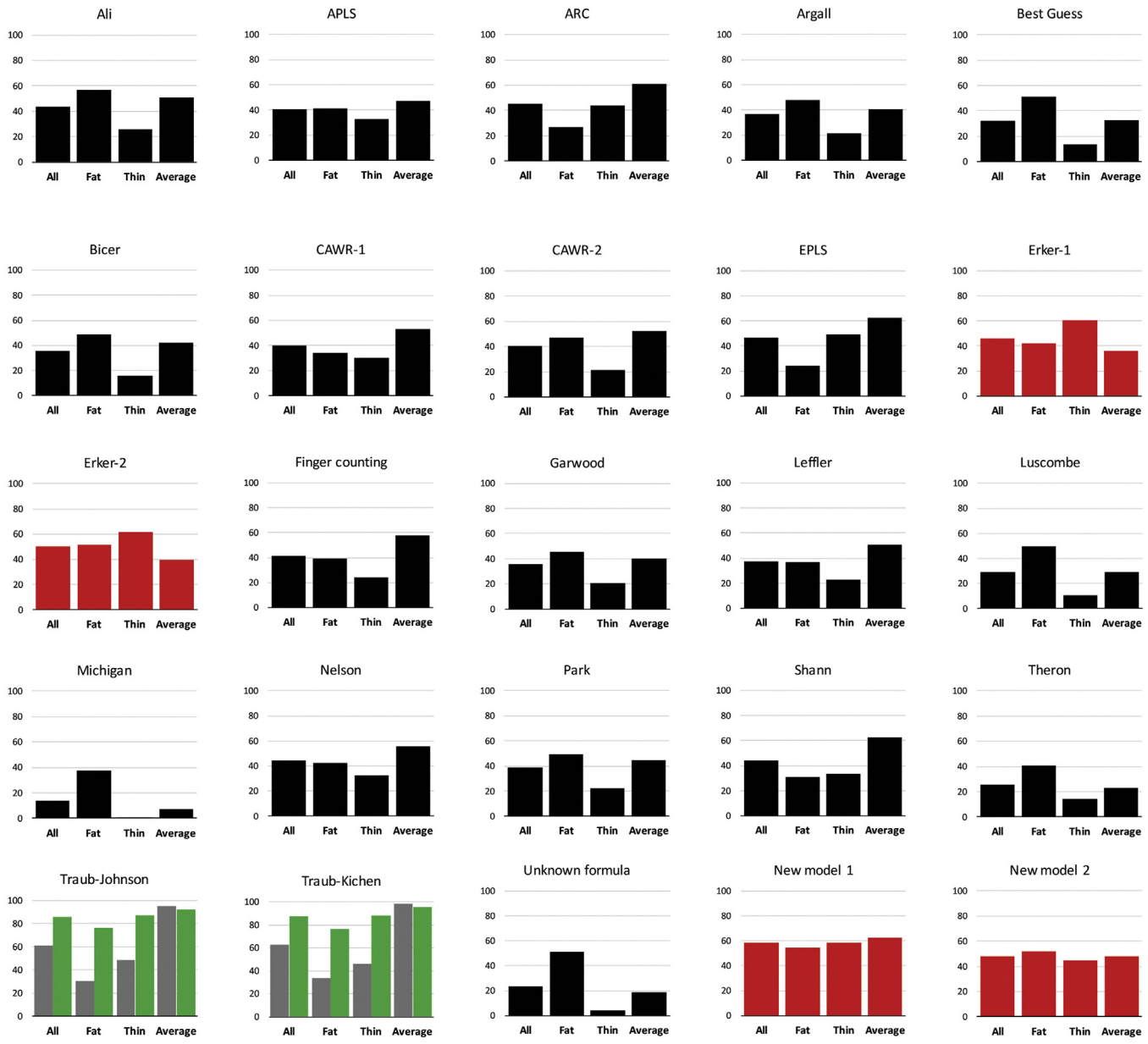


Fig. 3. (*Continued*)

types. These improvements allowed the formulas to achieve the predetermined acceptable outcome criteria. There was a slight, non-significant decrease in accuracy in children with an "average" body type in both these formulas.

Discussion

There are three broad points of interest arising from this study – the performance of the regular, unmodified aged-based formulas in this population, secondly, the performance of age-based, habitus-modified systems and, thirdly the performance of the length-based and length-based, habitus-modified formulas.

The performance of unmodified age-based weight-estimation formulas

Age-based formulas performed very poorly in this study. No formula achieved acceptable accuracy and there was a very high rate of critical errors: from 25% at best with the EPLS and ARC formulas to 70% with the Michigan formula. These results echo those from many other studies which have shown inadequate accuracy [23–25]. No formula has ever outperformed any length-based system (e.g. Broselow tape, PAWPER tape, Mercy method) in any study to date. Furthermore, as the poor precision of age formulas has repeatedly shown, no population can be homogeneous enough for them to be effective. Some formulas clearly meant for specific populations (e.g. the Park formulas for Korean children, the CAWR formulas for Hong Kong children, or the Theron formulas for Pacific Island children) have been shown to lack accuracy even in the populations from which they were derived.

New formulas are published regularly with no convincing evidence of their validity or adequate accuracy. As late as 2014, some researchers still believed that: "Age-based weight estimation formulas remain the only methods that have been validated in emergency situations, and that can help calculations of fluids, glucose, drug concentrations, etc., before the crucially ill child arrives" [26,27]. Neither part of this statement is true and this misconception may predispose children to avoidable medication errors. The

benefit of using formulas to allow advanced preparation (based on a weight that has a high probability of being inaccurate) must be more than offset by the detriments of the proven inaccuracy of the formulas and the proven inherent risk of calculation errors.

Most formulas were biased to overestimation of weight in this study population, given the relatively high prevalence of underweight children, as has been shown previously in developing countries like Malawi and Mali [28,29]. Other studies from populations with a high prevalence of overweight and obese children have shown the opposite, with a bias to underestimation [30]. The poor precision (wide limits of agreement) has been consistent across all populations in all studies, however.

The performance of age-based, habitus modified weight-estimation systems

Even though the age-based, habitus-modified formulas performed better than the unmodified formulas, they were still not acceptably accurate for clinical use because of the high number of critical errors. The Erker formulas, which showed great promise from the derivation study, the habitus-modified EPLS formula and the newly derived age-habitus formulas also performed poorly, although better than the unmodified formulas. They did not satisfy the predetermined acceptability outcome measures and were still far less accurate than previous reports on the Mercy method, the PAWPER tape or even the Broselow tape. The inherent biological variability of weight-for-age overwhelms the advantage of habitus modification, principally because length is not accounted for: a tall fat child will weigh more than a short fat child of the same age and habitus.

Assessment of habitus

The concept of habitus assessment by visual inspection has been proven in previous studies, but recent experiences with the PAWPER tape in obese populations have suggested that untrained individuals might not identify obesity well [31,32]. New methods of standardising habitus assessment, such as the use of figural reference images, may be needed in addition to formalised training to ensure that habitus assessment is optimised and standardised for whatever methodology it is used [21]. In this study, the habitus scores assigned by different raters were effective in discriminating between habitus-types and classifying children into significantly differentiated BMI strata. The ultimate proof that habitus assessment was accurate in this study was the dramatic effect that habitus-modification had on the length-based formulas in this study.

The performance of length-based formulas with and without habitus-modification

Ironically, the outcomes of the unmodified, very old length-based formulas (from 1980 and 1983, based on 1959 growth charts) were significantly superior to every one of the original and habitus-modified age-based formulas. The length formulas have not been studied much heretofore, perhaps because they are mathematically complex, but two previous studies have also shown better results when compared them with age-formulas [12,17]. The effect of applying a habitus-modification model to these formulas was astounding, producing a 40% relative improvement in accuracy and surpassing the specified minimum acceptable outcome criteria. The performance of both formulas was statistically similar and they both showed excellent consistency across the age and habitus subgroups. Both the formulas were inadvertently applied to children under one year of age (although with the length restriction applied for the Traub-Kichen formula), but the outcomes in this age-group were excellent and no different to that of older children. Since they are complex, these formulas would need to be incorporated into a computerised system or mobile device application for practical usefulness in emergencies, however.

The use of formulas in states of high cognitive stress

Even simple formulas cannot be used during emergencies without a high risk of error [33,34]. If $weight = agex2 + 8$ cannot be used safely, then $weight = 2.05 \times e^{0.02 \times length}$ plus habitus modification would be impossible. In emergency weight estimation, there needs to be a balance between the accuracy and ease-of-use of a method and strategies must be developed to achieve both these objectives. One such strategy would be a mobile phone application which could process the entered information to generate drug doses from length and body habitus without the user having to use cognitively demanding thought processes.

Limitations of this study

The patient population from this study might not be completely generalizable to other populations with a different prevalence of underweight and obesity, but it did contain a sizeable number of underweight, overweight and obese children. Although there might be slight differences in outcomes in a different population, there is a strong probability that the results will be similar – poor performance of all the age formulas and good performance of the length-based habitus-modified formulas.

Since the habitus score assessment depends on the rater's concept of "average" (and variations above and below average), the visual system of assigning a habitus score might be less accurate in a different population with a different prevalence of underweight or obese children. Whether different healthcare providers can accurately and reproducibly evaluate habitus in different environments, and what tools, such as reference images, might be required to achieve this still needs to be elucidated.

Conclusions

No unmodified age- or length based formula, of the more than 20 evaluated, reached the benchmark acceptable degree of accuracy, nor has any formula ever been shown to achieve acceptable accuracy in previous studies. The inescapable conclusion is that age-based formulas should not be used for weight estimation in children except *in extremis* and if no better method is available [35]. Furthermore, the use of habitus modification failed to improve the accuracy of the age-based formulas to an acceptable degree. The use of a visual assessment of habitus substantially improved the accuracy of the length-based formulas to the same or higher level of accuracy as other existing dual length- and habitus-based systems. Although promising, this needs further investigation, as does the methodology of habitus assessment. It proved successful in this study, but still needs further evaluation in different environments, by different observers, to ensure that standardisation of usage across populations is possible. Finally, formulas of any sort are at risk during periods of high cognitive stress and should not be used unless with the assistance of a proven electronic system (and not just a calculator!)

Conflicts of interest

We declare that we have no financial conflicts of interest. Although MW is the developer of the PAWPER tape, he derives no financial or commercial benefit from it.

Acknowledgement

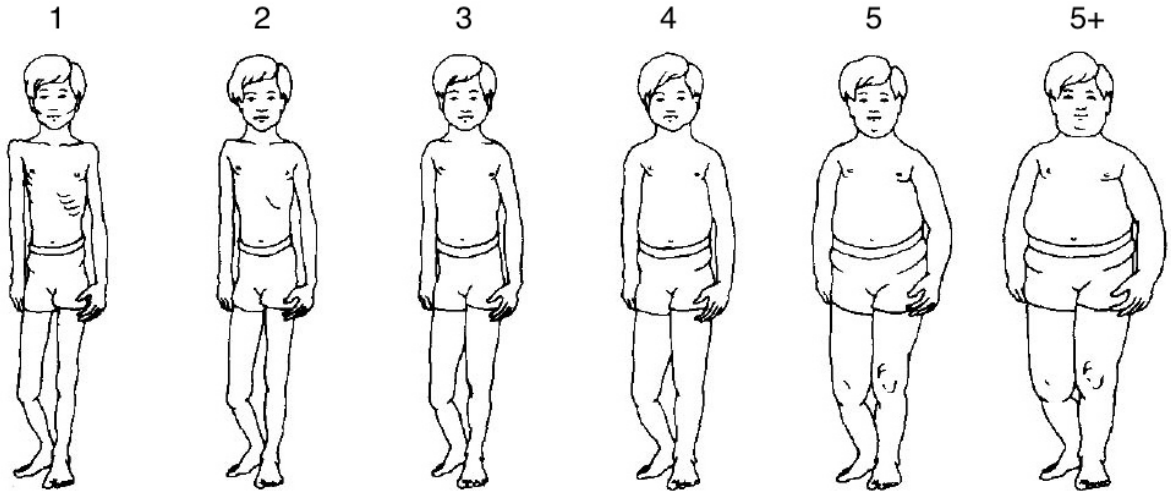
Nil.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.resuscitation.2017.05.018.

References

[1]. Young TP, Chen BG, Kim TY, Thorp AW, Brown L. Finger counting: an alternative method for estimating pediatric weights. Am J Emerg Med 2014;32(3):243–7.
[2]. Hegazy M, Taher E. Validating a new formula for weight estimation in pediatric cancer patients. Int Res J Med Med Sci 2013;1(1):34–9.
[3]. Leffler S, Hayes M. Analysis of parental estimates of children's weights in the ED. Ann Emerg Med 1997;30(2):167–70.
[4]. Batmanabane G, Kumar Jena P, Dikshit R. Using the Mercy Method for weight estimation in Indian children. Glob Pediatr Health 2015;(January–December):1–7.
[5]. Allison N, Norton I. Emergency weight estimation in Aboriginal and Torres Strait Islander children in the Northern Territory: are the current methods accurate? Emerg Med Austr 2014;26(5):487–93.
[6]. Wozniak R. The evaluation of potential weight-estimation methods in a primarily HIV positive cohort in Botswana for use in resource limited settings. The University of British Columbia; 2012.
[7]. Ackwerh R, Lehrian L, Nafiu OO. Assessing the accuracy of common pediatric age-based weight estimation formulae. Anesth Analg 2014;118(5):1027–33.
[8]. Bowen L, Zyambo M, Snell D, Kinnear J, Bould MD. Evaluation of the accuracy of common weight estimation formulae in a Zambian paediatric surgical population. Anaesthesia 2017;72(4):470–8.
[9]. Park J, Kwak YH, Kim do K, Jung JY, Lee JH, Jang HY, et al. A new age-based formula for estimating weight of Korean children. Resuscitation 2012;83(9):1129–34.
[10]. Britnell S, Koziol-McLain J. Weight estimation in paediatric resuscitation: a hefty issue in New Zealand. Emerg Med Austral 2015;27(3):251–6.
[11]. Dearlove CB, Dearlove O. Visual estimation of children's weights. Anaesthesia 1999;54(12):1228–9.
[12]. Black K, Barnett P, Wolfe R, Young S. Are methods used to estimate weight in children accurate? Emerg Med 2002;14(2):160–5.
[13]. Young KD, Korotzer NC. Weight estimation methods in children: a systematic review. Ann Emerg Med 2016;68(4):441–51, e10.
[14]. Garland JS, Kishaba RG, Nelson DB, Losek JD, Sobocinski KA. A rapid and accurate method of estimating body weight. Am J Emerg Med 1986;4(5):390–3.
[15]. Yamamoto LG, Inaba AS, Young LL, Anderson KM. Improving length-based weight estimates by adding a body habitus (obesity) icon. Am J Emerg Med 2009;27(7):810–5.
[16]. Wells M, Coovadia A, Kramer E, Goldstein L. The PAWPER tape: a new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. Resuscitation 2013;84(2):227–32.
[17]. Abdel-Rahman SM, Ridge AL. An improved pediatric weight estimation strategy. Open Med Dev J 2012;4:87–97.
[18]. Tanner D, Negaard A, Huang R, Evans N, Hennes H. A Prospective Evaluation of the Accuracy of Weight Estimation Using the Broselow Tape in Overweight and Obese Pediatric Patients in the Emergency Department. Pediatr Emerg Care 2016, http://dx.doi.org/10.1097/PEC.0000000000000894 [Epub ahead of print] PMID: 27741075.
[19]. Erker CG, Santamaria M, Moellmann M. Size does matter–age-related weight estimation in tall n' thin and tiny n' thick children and a new habitus-adapted alternative to the EPLS-formula. Resuscitation 2014;85(9):1174–8.
[20]. Wells M, Coovadia A, Kramer E, Goldstein L. The PAWPER tape: A new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. Resuscitation 2013;84(2):227–32.
[21]. Georgoulas V, Wells M. The PAWPER tape and the Mercy Method outperform other methods of weight estimation in children in South Africa. South Afr Med J 2016;106(9):933–9.
[22]. Stewart D. Accuracy of the Broselow tape for estimating paediatric weight in two Australian Emergency Departments. University of Sydney; 2009.
[23]. Marlow R, Lo D, Walton L. Accurate paediatric weight estimation by age: mission impossible? Arch Dis Child 2011;96(Suppl. 1):A1–2.
[24]. Cattermole GN, Leung MP, So HK, Mak PS, Graham CA, Rainer TH. Age-based formulae to estimate children's weight in the emergency department. Emerg Med J 2011;28(5):390–6.
[25]. Cattermole GN, Graham CA, Rainer TH. Age-based weight prediction is imprecise. Emerg Med J 2011;29(1):82–3.
[26]. Skrobo D, Kelleher G. CORKSCREW 2013 CORK study of children's realistic estimation of weight. Emerg Med J 2015;32(1):32–5.
[27]. L. Janeckzo. Two easy tools estimate children's body weight without scales 2014. Available from: http://www.consultant360.com/story/two-easy-tools-estimate-childrens-body-weight-without-scales. [Acessed 27, November 2015].
[28]. Pollock L, Molyneux E, Jefferis O, Bunn J. Estimating children's weights: a developing country perspective. Arch Dis Childhood 2007;92(10):938–9.
[29]. Dicko A, Alhousseini ML, Sidibé B, Traoré M, Abdel-Rahman SM. Evaluation of the Mercy weight estimation method in Ouelessebougou, Mali. BMC Public Health 2014;14:270.
[30]. Luscombe M, Owens B. Weight estimation in resuscitation: is the current formula still valid? Arch Dis Childhood 2007;92(5):412–5.
[31]. Chavez H, Peterson R, Lo K, Arel M. Weight estimation in an inner city Pediatric Emergency Department: the effect of obesity. Am J Emerg Med 2015;33(10):1364–7.
[32]. Garcia CM, Meltzer JA, Chan KN, Cunningham SJ. A validation study of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape-a new weight estimation tool. J Pediatr 2015;167(1):173–7, e1.
[33]. Lancaster L. Mission impossible—minimising the terror of paediatric resuscitation for staff in the ED. Austral Emerg Nurs J 2002;5(3):16–20.
[34]. Marikar D, Varshneya K, Wahid A, Apakama O. Just too many things to remember? A survey of paediatric trainees' recall of Advanced Paediatric Life Support (APLS) weight estimation formulae. Arch Dis Childhood 2013;98(11):921.
[35]. S. Carley. APLS weight estimation -don't do it (well almost never). 2013. Available from: http://stemlynsblog.org/apls-estimation-formulas-do-not-safely-predict-weight-in-uk-children-st-emlyns/-comments. [Acessed 27, November 2015].



Supplementary figure Figural reference images for habitus score assessment. Modified Collins figures were developed and validated for habitus assessment, with an accuracy shown to be equal to that of the gestalt visual inspection system [21]. Reference images can be used for training or to increase the objectivity of habitus assessment in populations where the “average” child has a habitus significantly above or below the 50th centile of weight-for-length.

5.3 Epilogue

5.3.1 Major findings

Age-based formulas are commonly researched, taught and used in simulated and real clinical practice despite strong evidence that they are not – and cannot – be accurate [4, 16]. This study, the most extensive of its kind to date, evaluated the performance of every existing published age-based weight estimation formula in a large study. It was also the first study to evaluate the only published age-based, habitus-modified formulas and to have evaluated length-based, habitus modified formulas.

The two major findings were quite simple: no age-based formula (including the new age- and habitus-based formulas) was accurate; while the length- and habitus-based formulas were extremely accurate at estimating weight. This strongly suggested that age is not as good a biological indicator of weight as is length.

The differences in outcomes between age formulas and methods such as the Mercy method and the PAWPER tape suggest that age-based weight estimation is simply too inaccurate to be used. The fact that habitus-modification is extremely useful with length-based systems, but not effective for age-based weight estimation, lends further support to the newer weight estimation systems. The length-based formulas with habitus modification did not achieve a level of accuracy that suggested that they should be investigated in preference to the newer dual length- and habitus-based methods. The complexity of the formulas also makes them less usable during emergency care than the simpler PAWPER tape.

5.3.2 This study in the context of the thesis

Although the evidence against the accuracy of age-based weight estimation is convincing, there are some additional facts that need to be considered. There is no evidence about how often age is available to emergency care providers who use age formulas, or how often an incorrect age is used. There is evidence that healthcare providers are not able to estimate age accurately and not able to apply formulas during emergencies [8, 17]. All this evidence supports the use of other methods of weight estimation and places age formulas firmly at the bottom of the hierarchy.

Interestingly, there is some preliminary evidence that age-based formulas may be useful for estimating IBW in children, rather than TBW or actual weight, which may be useful in drug dosing for obese children [3, 18]. This potential application will be discussed in Chapter 7.

After the publication of this study, a new age- and habitus-based method was published by Appelbaum and colleagues [19]. It makes use of age, gender and a visual assessment of habitus to estimate weight. The validation study in a large database showed very promising results, but these will need to be verified in a prospective study. The major problem with this initial study was that each child's habitus assessment was retrospectively derived from their BMI-for-age centile in the database. This created a strong statistical linkage to length, which might not hold true when prospective data is used. Nonetheless, the confirmation of the value of habitus assessment provided by this study was important. The Appelbaum study also used the same habitus category structure as the PAWPER XL tape, which the authors found to be statistically consistent with the BMI-for-age centiles used in the PAWPER system.

5.3.3 Link to the next article

The meta-analysis in Chapter 4 and the focussed evaluation of age-based formulas in this chapter have provided strong evidence that these formulas should not be used in clinical practice unless new and compelling evidence to the contrary emerges. This was important evidence to ascertain, given the common usage of formulas.

As part of this thesis, it was also important to evaluate one of the most important weight estimation systems in existence today: the Broselow tape. The meta-analysis in Chapter 4 showed the tape to be more accurate than age-formulas, but not as accurate as the length- and habitus-based systems. While this tape is often considered to be the "gold standard" for weight estimation, many authors have queried its accuracy in the last decade [20, 21]. For this reason, it was crucial to systematically review the evidence on the accuracy of the Broselow tape as well as the evidence for the drug doses and equipment sizes printed on the tape itself. This

feature is what accounts for its worldwide popularity, but this might not compensate for its other shortcomings.

5.3.4 References

[1] Wells M, Goldstein L, Bentley A. The accuracy of emergency weight estimation systems in children - a systematic review and meta-analysis. International Journal of Emergency Medicine. 2017;10(29):1-43. https://doi.org/10.1186/s12245-017-0156-5.

[2] Wells M, Goldstein L, Bentley A. A systematic review and meta-analysis of the accuracy of weight estimation systems used in paediatric emergency care in developing countries. African Journal of Emergency Medicine. 2017;7:S36-S54.

[3] Young KD, Korotzer NC. Weight estimation methods in children: a systematic review. Annals of Emergency Medicine. 2016;68(4):441-451 e10.

[4] Marlow R, Lo D, Walton L. Accurate paediatric weight estimation by age: mission impossible? Archives of Disease in Childhood. 2011;96(Suppl 1):A1-A2.

[5] Wells M, Barnes L, Vincent-Lambert C. Paediatric weight estimation practices of advanced life support providers in Johannesburg, South Africa. African Journal of Emergency Medicine. 2018;8(2):51-54.

[6] Advanced Life Support Group. Advanced Paediatric Life Support: The Practical Approach (APLS) 6th Edition. New York: BMJ Books - Publisher: John Wiley & Sons (Wiley-Blackwell); 2016.

[7] Kaufmann J, Laschat M, Wappler F. Medication errors in pediatric emergencies: a systematic analysis. Deutsches Ärzteblatt International. 2012;109(38):609-16.

[8] Marlow RD, Wood DLB, Lyttle MD. Comparing the usability of paediatric weight estimation methods: a simulation study. Archives of Disease in Childhood. 2018.

[9] Lancaster L. Mission impossible — minimising the terror of paediatric resuscitation for staff in the ED. Australasian Emergency Nursing Journal. 2002;5(3):16-20.

[10] Marikar D, Varshneya K, Wahid A, Apakama O. Just too many things to remember? A survey of paediatric trainees' recall of Advanced Paediatric Life Support (APLS) weight estimation formulae. Archives of Disease in Childhood. 2013;98(11):921.

[11] Deboer S, Seaver M, Broselow J. Color coding to reduce errors. American Journal of Nursing. 2005;105(8):68-71.

[12] Erker CG, Santamaria M, Moellmann M. Size does matter--age-related weight estimation in "tall n' thin" and "tiny n' thick" children and a new habitus-adapted alternative to the EPLS-formula. Resuscitation. 2014;85(9):1174-1178.

[13] Traub SL, Kichen L. Estimating ideal body mass in children. American Journal of Hospital Pharmacy. 1983;40(1):107-10.

[14] Traub SL, Johnson CE. Comparison of methods of estimating creatinine clearance in children. American Journal of Hospital Pharmacy. 1980;37(2):195-201.

[15] Goldman J, Ojoo A, Abdel-Rahman S. Challenges in pediatric oral dosing. In: MacLeod S, Hill S, Koren G, Rane A, editors. Optimizing treatment for children in the developing world. Cham: Springer International Publishing; 2015. p. 33-44.

[16] Cattermole GN, Graham CA, Rainer TH. Age-based weight prediction is imprecise. Emergency Medicine Journal. 2011;29(1):82-83.

[17] Williams B, Boyle M, O'Meara P. Can undergraduate paramedic and nursing students accurately estimate patient age and weight? Prehospital and Disaster Medicine. 2010;25(2):171-177.

[18] Carasco CF, Fletcher P, Maconochie I. Review of commonly used age-based weight estimates for paediatric drug dosing in relation to the pharmacokinetic properties of resuscitation drugs. British Journal of Clinical Pharmacology. 2016;81(5):849-56.

[19] Appelbaum N, Clarke J, Maconochie I, Darzi A. Paediatric weight estimation by age in the digital era: optimising a necessary evil. Resuscitation. 2018;122:29-35.

[20] Nieman CT, Manacci CF, Super DM, Mancuso C, Fallon WF, Jr. Use of the Broselow tape may result in the underresuscitation of children. Academic Emergency Medicine. 2006;13(10):1011-9.

[21] Clark MC, Lewis RJ, Fleischman RJ, Ogunniyi AA, Patel DS, Donaldson RI. Accuracy of the Broselow Tape in South Sudan, "The Hungriest Place on Earth". Academic Emergency Medicine. 2016;23(1):21-28.

Chapter 6

The accuracy of the Broselow tape as a weight estimation tool and a drug-dosing guide.

6.1 Prologue

6.1.1 Background

The meta-analysis in Chapter 4 showed that the Broselow tape was more accurate than age-based formulas but not as accurate as the dual length- and habitus-based systems. The Broselow tape is one of the most widely-used weight estimation systems across the world for two reasons: it is perceived to be very accurate and users favour the drug and equipment guidelines printed on the tape [1]. The evidence supporting both of these factors has been called into question in recent studies, however [2, 3]. The study presented in this chapter is the first study to have critically evaluated the functioning of the Broselow tape both as a weight estimation tool and a drug dosing guide.

One of the considerations in the development of the PAWPER system was whether it would be useful to have drug and equipment information printed on a new version of the tape. There are now a number of tapes similar to the Broselow tape which contain pre-calculated drug dosing information in a similar format [4]. Some authors have considered this information to be incomplete, however, as additional reference material is still required to eliminate the need for calculations (needed for the dilution and preparation of drugs) and minimise cognitive load [5]. A comprehensive systematic review had never been performed on the Broselow tape's role in weight estimation or its role as a resuscitation aid. The information gathered in this review would be necessary to elucidate the potential benefits of having limited information on the tape, which could guide the future development of the PAWPER system.

Other available evidence from studies on the Broselow tape have shown that even simple tape devices can be vulnerable to error when used in emergencies by individuals without sufficient training [6-8]. This assessment of human factor errors was noteworthy as it emphasised the need to keep weight estimation systems simple

and to provide users with sufficient training with the devices. This is covered in more detail in Chapter 10.

6.1.2 Purpose of the study

The two main purposes of this study were to critically evaluate the Broselow tape's role as a weight estimation device and to evaluate the evidence of the efficacy of its role in guiding emergency drug dosing.

There are concerns about the accuracy of the Broselow tape in both high-income and low- and middle-income populations [9-12]. The first purpose of this study was to review the accuacy of the Broselow tape as a weight estimation instrument. Although the meta-analysis in Chapter 4 showed that the tape was not as accurate as the dual length- and habitus-based systems, it was important to focus in more depth on the Broselow tape, given its wide usage and standing as the "gold standard" of weight estimation. Recent evidence from low- and middle-income countries showing that the current edition of the Broselow tape, the 2011A, was less accurate than previous editions was also of concern. This study also reviewed evidence on the changes in the tape since its inception and explored the effects of these changes in populations with differing proportions of underweight and obese children.

The second purpose of this paper was to review the evidence behind the Broselow tape as a drug dosing guide. This would establish whether there is inherent value in the printed drug doses on a tape, in order to decide whether this was the most appropriate design to incorporate into the newer weight estimation devices.

This study provided data to contribute to both Objectives 1 and 2 (to describe the accuracy of commonly-used weight estimation systems in both high-income and low- and middle-income countries through the systematic review of previous studies and meta-analysis of data obtained from these studies *and* to review the evidence supporting the inclusion of drug dosing information on length-based weight-estimation tapes, specifically the Broselow tape).

6.2 Published paper

Wells M, Goldstein L, Bentley A, Basnett S, Monteith I. The accuracy of the Broselow tape as a weight estimation tool and a drug-dosing guide - a systematic review and meta-analysis. Resuscitation. 2017;121: 9-17.

The PDF of the published article is included. The online-only supplementary material is appended after the PDF. The tables or figures have been retained in the formatting in which they are archived in the journal website.

Resuscitation 121 (2017) 9–33

Contents lists available at ScienceDirect

Resuscitation

journal homepage: www.elsevier.com/locate/resuscitation

Review

The accuracy of the Broselow tape as a weight estimation tool and a drug-dosing guide – A systematic review and meta-analysis

Mike Wells*, Lara Nicole Goldstein, Alison Bentley, Sian Basnett, Iain Monteith

Division of Emergency Medicine, Faculty of Health Sciences, University of the Witwatersrand, 7 York Road, Parktown, Johannesburg, South Africa

ARTICLE INFO

Article history:
Received 29 May 2017
Received in revised form 31 August 2017
Accepted 23 September 2017

Keywords:
Broselow tape
Paediatric weight estimation
PAWPER tape
Mercy method
Resuscitation aid

ABSTRACT

Aims: The Broselow tape is widely used as a weight-estimation device and drug-dosing guide aid, but concerns about its accuracy and its efficacy have emerged in the last decade. The aim of this study was to systematically review the literature to analyse the accuracy of the Broselow tape as a weight estimation device and review evidence of its utility as a drug-dosing guide.
Methods: This was a MOOSE-driven systematic review and meta-analysis, which focused on studies evaluating the accuracy of the Broselow tape and studies reviewing its use as a drug-dosing aid.
Main results: The tape has undergone substantial changes over the years, but there was no evidence to show that the changes have improved weight-estimation performance. The weight-estimation accuracy of the tape was suboptimal in all populations, with just over 50% of children receiving an estimation within 10% of their actual weight. The overestimation of weight in low- and middle-income countries was often extreme. This indicated a significant potential for potentially harmful medication errors. The limited available evidence on the value of the tape as a drug-dosing guide indicated that the tape was frequently used incorrectly and contained insufficient information to function without additional resources.
Conclusions: The Broselow tape lacked sufficient accuracy as a weight estimation and drug-dosing tool when compared to other available techniques. In addition, the Broselow tape contains insufficient drug-dosing information to function as a complete resuscitation aid without additional material. The frequent rate of incorrect usage of the tape indicated that appropriate training with the tape is mandatory to reduce errors.

© 2017 Elsevier B.V. All rights reserved.

1 Introduction

The Broselow tape has not been the subject of systematic review or meta-analysis previously. This review is important because a knowledge and understanding of the tape, its different functions, its strengths and its weaknesses will assist its potential users to decide whether it is the best weight estimation system and drug-dosing guide for their population and their particular needs. It will also assist decision-makers in the Emergency Department and Emergency Medical Services to develop appropriate evidence-based approaches to emergency weight estimation and emergency drug dosing in children.

The Broselow tape was developed primarily as a tool to facilitate the delivery of rapid, appropriate and safe treatment during emergencies, but most of the research published on the tape has focused on the accuracy of the weight estimation [1–3]. The Broselow tape has not been without controversy: some experts have suggested that the Broselow tape, because of its underestimation of weight in high-income countries, would lead to the "under-resuscitation of children" and that it might even no longer be relevant [4,5]. At the other end of the habitus spectrum, there is also evidence that the tape, especially in its latest edition, substantially overestimates the weight of children in populations from low- and middle-income countries, often to a potentially dangerous degree [6–8]. Since the Broselow tape is still widely used across the world as a weight-estimation tool,an understanding of its accuracy, especially when

Abbreviations: 2D, two-dimensional; 3D, three-dimensional; APLS, Advanced Paediatric Life Support formula; ARC, Australian Resuscitation Council formula; BG, Best Guess formula; BMI, body mass index (kg/m^2); BT, Broselow tape; DWEM, devised weight-estimating method; ED, emergency department; EMS, emergency medical services; EPLS, European life support formula; FE, fixed effects; HCP, healthcare provider; IBW, ideal body weight; IQR, interquartile range; IV, inverse variance; MAC, mid-arm circumference; MPE, mean percentage error; MT, mercy tape; NCHS, National Centre for Health Statistics; PAWPER, paediatric advanced weight prediction in the emergency room; PW10, percentage of estimates within 10% of actual weight (%); PW20, percentage of estimates within 20% of actual weight (%); RE, random effects; TBW, total body weight; TJ, Traub-Johnson formula; TK, Traub-Kichen formula.

* Corresponding author.
E-mail addresses: mike.wells@emergencymedicine.co.za (M. Wells), drg666@gmail.com (L.N. Goldstein), Allison.bentley@wits.ac.za (A. Bentley), sibazz26@gmail.com (S. Basnett), montyza@gmail.com (I. Monteith).

https://doi.org/10.1016/j.resuscitation.2017.09.026
0300-9572/© 2017 Elsevier B.V. All rights reserved.

compared to other, newer weight estimation systems, would be useful.

The Broselow tape's effectiveness as a resuscitation aid to provide critical drug-dosing information has also been questioned [9,10]. This raises interesting new uncertainties. This is significant because some experts, and advanced life support courses, advocate the use of "a body length tape with pre-calculated doses", which may, in fact, not be compatible with best evidence-based practice [11].

Therefore, to minimise the potential for medication errors during the emergency management of children, it would be valuable to conduct a careful "re-examination of the accuracy of the Broselow tape" [12].

The objectives of this review were:

- To review the available evidence on the accuracy of the Broselow tape as a weight-estimation instrument in different populations and perform a *meta*-analysis on this data.
- To review the available evidence on the efficacy of the tape as a drug-dosing guide during the management of paediatric emergencies.

2 Methods

2.1 Systematic review and meta-analysis

This study followed the guidelines for meta-analyses and systematic reviews of observational studies (MOOSE-guidelines).

2.2 Search strategy

Online databases (MEDLINE, SCOPUS, Science Direct and Google) were searched for appropriate studies, published between January 1986 and January 2017, using the search terms: "paediatric weight estimation", "weight estimation children", "resuscitation aid" and "Broselow tape". Citations from reference lists, abstracts, conference presentations and unpublished material were also scrutinised to identify articles for potential inclusion.

2.3 Study selection and eligibility criteria

All studies that appraised, reviewed or discussed the Broselow tape were separately considered for inclusion by two investigators (MW and LG). There were two areas of interest: the accuracy of the Broselow tape as a weight estimation system and the Broselow tape as a medication-dosing guide (see Fig. 1). Each study was graded for quality of evidence using a modified Newcastle-Ottawa scale (see Supplementary Table 1) and for risk of bias. Evaluated studies were downgraded if there were significant methodological weaknesses (e.g. length data converted to Broselow tape weight estimate), if data presentation was incomplete and if performance outcome data was not appropriately converted to percentage or logarithmic format. Studies were excluded from the meta-analysis if they did not contain appropriate data describing bias and precision (mean percentage error plus any appropriate measure of variance) *or* data describing accuracy (percentages of weight estimations accurate to within 10% of actual weight).

2.4 Data abstraction and analysis

Mean percentage error (MPE) with the 95% limits of this error were used to evaluate the bias and precision of the Broselow tape. The percentage of weight estimations within 10% of actual weight (PW10) was used to evaluate the accuracy of the Broselow tape. Standard statistics for meta-analysis of method-comparison studies were used, but the included studies showed a significant amount of within-study variance as well as between-study variance that needed to be considered [13]. Two methods of representing the pooled data were employed: a fixed effects model weighted by inverse variance (method 1); and a random effects model (method 2). The sizeable variance between studies, the inclusion of several large database studies with substantial within-study variance and the heterogeneity between populations meant that the random effects models were preferred for most analyses. Where more than one study contained data on the Broselow tape as well as other weight-estimation methodologies, these were directly compared using accuracy outcomes in a random-effects, inverse variance-weighted model.

2.5 Risk of bias within and across studies

Many of the studies were at risk of performance bias because the measuring personnel were not blinded during the execution of weight estimation (e.g. the use of the Broselow tape) and there was similarly a risk of detection bias, as the outcome assessors were not blinded, but, since most of the data were in simple, objective, numerical form, this lack of blinding was believed unlikely to cause significant bias. Reporting bias (or publication bias) was minimised by including all available methodologically sound studies in the review, published or not. Methodological causes of potential bias were common (e.g. the Broselow tape was not actually used in many studies, but weight-estimates were instead generated from length data), but these were individually assessed and rated according to the level of risk of systematic bias. Evaluation of pooled data with and without studies of uncertain risk yielded similar outcomes. Further evaluation of data with and without the inclusion of the very large, retrospective database studies also produced similar outcomes, which suggested a negligible bias caused by these studies. Studies with high risk of bias were excluded from the meta-analysis, however. Table 1 provides a descriptive summary of the studies included in this review. The raw data from the included studies is shown in Supplementary Table 2.

The quality of the evaluated studies was rated according to the prescribed MOOSE criteria, including rationale for the selection of participants, assessment of confounding factors, assessment of study quality, including blinding of quality assessors, assessment of heterogeneity, description of statistical methods and provision of appropriate tables and graphics.

2.6 Sensitivity analysis

Since the outcomes between study populations from high-income and low- and middle-income countries were markedly different, these studies were also evaluated within different categories. While the prevalence of obesity may also be significant in low- and middle-income countries (e.g. 24% in Botswana, 14% in Sudan), the real difference between high-income and low- and middle-income countries may be the prevalence of underweight children rather than obesity only [14]. Three categories were considered for the subgroup analysis, based on the prevalence of underweight children in that country or region: Category I – low (<2%) prevalence of underweight children (high-income countries); Category II – medium (2–8%) prevalence of underweight (generally middle-income countries e.g. Mexico, Thailand, Iran, parts of South Africa); and Category III – high (>8%) prevalence of underweight (generally low- and middle-income countries e.g. India, most African countries).

2.7 Subgroup analysis

Wherever possible, subgroup data, by age or weight, that were presented in each study were included and evaluated in the overall meta-analysis strategy.

Table 1
Studies included in the review. Basic information is shown, including an assessment of risk of bias and the Newcastle-Ottawa Scale star-rating of study quality (maximum score of 10 stars). The first part of the table (1a) shows studies with weight estimation data that were evaluated for inclusion in the quantitative analysis of weight-estimation accuracy. The second part of the table (1b) shows the studies included for the assessment of the accuracy of the Broselow tape as a resuscitation and drug-dosing aid or for qualitative data.

(a)

Author, date	Study size (N)	Country	Design	Patient ages	Estimation techniques evaluated	Included in meta-analysis	Main results, Newcastle-Ottawa rating, Risk of Bias assessment
Lubitz 1988 [2]	937	USA	P	0 to 12yrs	Broselow tape	Yes	**Main Results:** This was the original study of the Broselow tape. The Broselow tape was better than healthcare provider guesses and had similar accuracy to DWEM. Accuracy of Broselow tape fell off sharply in children >25 kg. The authors recommended that an assessment of body habitus should be considered in children >25 kg **Weaknesses:** No formal, prospective comparison with other methodologies or indication of desirable accuracy. ★★★★★★ Risk of bias: low
Hughes 1990 [3]	139	UK	P	0 to 10yrs	Broselow tape, healthcare provider guesses	No	**Main Results:** This was the first validation study of the Broselow tape in the UK. The Broselow tape performed substantially better than nurses' guesses. **Weaknesses:** Small sample size. Limited statistics with incomplete data. ★★★★ Risk of bias: high (incomplete data and/or statistical analysis)
Dearlove 1999 [15]	50	UK	P	1 to 16yrs	Parental estimates, Broselow tape, EPLS, Argall	Yes	**Main Results:** The Broselow tape performed best, far better than parental estimates and age-based formulas. A target of 10% accuracy was considered appropriate for children. **Weaknesses:** Limited statistical analysis and very small sample size. ★★★★ Risk of bias: low
Kun 2000 [16]	909	Hong Kong	P	0 to 12yrs	Broselow tape	Yes	**Main Results:** The Broselow tape was most accurate in children 10–25 kg. Adjustment for habitus recommended by authors. Objective assessment of the accuracy of Broselow tape outside of the 10–25 kg range showed that it was actually moderate and not as good as the authors suggest. **Weaknesses:** Incorrect interpretation of some statistics. Broselow tape version not reported. ★★★★ Risk of bias: low
Black 2002 [17]	495	Australia	P	0 to 18yrs	EPLS, Broselow tape, DWEM, Oakley, TJ, TK	Yes	**Main Results:** The Broselow tape and DWEM were more accurate than formulas. These methods should be used if weighing not possible. The EPLS was the worst performer but there was poor accuracy of all systems. Good reproducibility of assessment of body habitus for the DWEM. **Weaknesses:** Limited and incomplete statistics. Broselow tape version not reported. ★★★★ Risk of bias: low
Hofer 2002 [18]	585	Switzerland	R	6mo to 11yrs	Broselow tape	Yes	**Main Results:** The Broselow tape was moderately accurate but underestimated weight in older children. Nearly 25% of sample excluded because they were too tall for the tape. **Weaknesses:** Broselow tape not actually used and version not reported. Incomplete reporting of data. ★★★★ Risk of bias: low
Argall 2003 [19]	300	UK	P	1 to 10yrs	EPLS, Broselow tape	No	**Main Results:** Both methods performed poorly and worsened with increasing age. Difficult to draw any conclusions from this study because of incomplete date presentation, but the Broselow tape was marginally better than the formula. The authors suggested that methods of weight estimation are not keeping up with increasing obesity. **Weaknesses:** Broselow tape version not reported. Incomplete statistical reporting made interpretation difficult. ★★★★ Risk of bias: unknown
Potier 2003 [20]	–	–	–	–	EPLS, Broselow tape	No	**Main Results:** this was a non-systematic mini- PICO analysis. The EPLS "may" be losing accuracy with increasing obesity; Broselow tape "may indeed be" more accurate. **Weaknesses:** Limited qualitative-only evaluation. No comment on acceptable degree of accuracy. ★★ Risk of bias: high
Theron 2005 [20]	909	New Zealand	P	1 to 10yrs	EPLS, Shann, Leffler, Oakley, Broselow tape	Yes	**Main Results:** Formulas and the Broselow tape underestimated the weight of Pacific Island and Maori children. The Broselow tape was the best performer of the systems tested, despite the fact that the authors reported the contrary. **Weaknesses:** Broselow tape version not reported. Limited and incomplete statistics. ★★★★ Risk of bias: low

Table 1 (*Continued*)

(a)

Author, date	Study size (N)	Country	Design	Patient ages	Estimation techniques evaluated	Included in meta-analysis	Main results, Newcastle-Ottawa rating, Risk of Bias assessment
Nieman 2006 [5]	7813	USA	R	0 to 12yrs	Broselow tape	Yes	**Main Results:** One third of children had inaccurate weight estimations. Surprisingly, the 1998 tape performed better than the 2002 tape. The authors suggested that a measurement of obesity (e.g. mid-arm circumference) should be added to Broselow tape to increase accuracy. Broselow tape version 1998 and 2002A used. The study population BMI was 16.8 kg/m^2. Good statistics. **Weaknesses:** Broselow tape not actually used. Nearly 7% of sample excluded because they were too tall for the tape. ★★★★★★★ Risk of bias: low
Varghese 2006 [21]	500	India (Cat III)	P	1 to 12yrs	Argall, EPLS, Nelson, Broselow tape	No	**Main Results:** Formulas overestimated weight in this low- and middle-income country study. The Broselow tape was the most accurate. More than half the study population was under 6 months of age and only 8% >5yrs. **Weaknesses:** Limited statistical reporting and data analysis made findings difficult to interpret. Age-range of sample very skewed to younger children. ★★★ Risk of bias: high
Du Bois 2007 [22]	400	USA	P	0 to 19yrs	Broselow tape, DWEM	Yes	**Main Results:** The DWEM performed better than the Broselow tape, but both systems underestimated weight, especially in children >20 kg. There was good inter-rater reliability for habitus assessment. Although the performance of DWEM was better than Broselow tape, both systems performed poorly. **Weaknesses:** Limited data and statistics. Broselow tape version not reported. Statistically inappropriate accuracy target suggested (bias confused with accuracy). ★★★★★★ Risk of bias: low
Hashikawa 2007 [23]	1207	USA	R	0 to 12yrs	Broselow tape	Yes	**Main Results:** Approximately 60% accuracy of colour zone assignment (i.e. accurate drug dosing). Weight was underestimated in obese and older children. The rising prevalence of obesity blamed for poor performance. Broselow tape version 2002B. The average study population BMI was 17 kg/m^2. **Weaknesses:** Broselow tape not actually used. Assessment of correct zone assignment only, weight not measured. Limited and incomplete statistics. Risk of bias: low
Im 2007 [24]	454	Korea	P	–	Broselow tape	No	**Main Results:** The Broselow tape was accurate in children of normal weight-for-length. Only children falling within 3rd to 97th weight-for-height centiles were included. Very young study population. Broselow tape only recommended by authors for "normal-growth" children <20 kg in weight and <120 cm in height. **Weaknesses:** poor methodology and limited statistics. Conclusions not justified or supported by findings. ★★★ Risk of bias: high
Jang 2007 [25]	665	Korea	R	0 to 12yrs	Broselow tape	Yes	**Main Results:** The Broselow tape was reasonably accurate in this population, but less so in children >25 kg. There was an overall underestimation of weight and the performance was not very good, although on par with most other studies. **Weaknesses:** Broselow tape not actually used and version not reported. Limited and incomplete statistics. ★★★★ Risk of bias: low
Krieser 2007 [26]	410	Australia	P	1 to 10yrs	Parental estimates, Broselow tape, BG, Argall, EPLS	Yes	**Main Results:** Parental estimates performed best, followed by the Broselow tape. Only 11% of parents could not provide an estimate. Formulas performed much worse than other methods. Study population BMI 17.1 kg/m^2. **Weaknesses:** Multiple papers on same data. Broselow tape not actually used. Broselow tape version not reported. Incomplete statistical analysis. ★★★★★ Risk of bias: low
Patel 2007 [27]	360	USA	P	1mo to 10yrs	Parental estimates, EPLS, Broselow tape	No	**Main Results:** Abstract. Parental estimates performed best, then Broselow tape, then EPLS. Conclusions based on correlation and rudimentary statistical analysis**. Weaknesses:** Limited data and statistics presented. Findings unreliable and hard to interpret. ★★★ Risk of bias: high

Table 1 (*Continued*)

(a)

Author, date	Study size (N)	Country	Design	Patient ages	Estimation techniques evaluated	Included in meta-analysis	Main results, Newcastle-Ottawa rating, Risk of Bias assessment
Zikos 2007 [28]	360	USA	P	1mo to 10yrs	Parental estimates, EPLS, Broselow tape	No	**Main Results:** Virtually identical abstract and data to Patel 2007. ★★★ Risk of bias: high
Ramarajan 2008 [29]	548	India (Cat III)	P	0 to 12yrs	Broselow tape	Yes	**Main Results:** The Broselow tape overestimated weight by >10% in Indian children over 10 kg. A correction factor was developed, but not validated. One of the earliest studies to show overestimation of weight by Broselow tape. **Weaknesses:** Broselow tape version not reported. Incomplete statistical analysis and data presentation. ★★★★★★★ Risk of bias: low
Zink 2008 [30]	127	USA	P	0 to 17yrs	Parental estimates, healthcare provider guesses, DWEM, Broselow tape	No	**Main Results:** The authors claimed that the Broselow tape and DWEM were the least accurate methods. No conclusions could be drawn from this study because of the defective methodology. **Weaknesses:** The data appears to favour healthcare provider and parent guesses, but the statistical methodology was flawed. ★★ Risk of bias: high
Anstett 2009 [31]	545	Ireland	R	–	Broselow tape	Yes	**Main Results:** Abstract. The Broselow tape was often inaccurate and tended to underestimate weight, but actually performed better in this study than in many other studies. **Weaknesses:** Broselow tape version not reported. Broselow tape not actually used. ★★★★★★ Risk of bias: low
Cattamanchi 2009 [32]	15000	India (Cat III)	P	2mo to 12yrs	Broselow tape	Yes	**Main Results:** Abstract. Very large prospective study. The Broselow tape performed well, especially in children <10 kg but underestimated all others, especially in children >18 kg. The authors recommended a new version of Broselow tape for Indian children because of underestimation of weight. **Weaknesses:** Broselow tape version not reported. Incomplete statistical analysis and some misinterpretation of data. ★★★★★★★ Risk of bias: low
Cattermole 2009 [33]	1368	Hong Kong	P	1 to 12yrs	MAC formula, Broselow tape, foot length	No	**Main Results:** Abstract. Estimates of weight can be based on mid-arm circumference. A special colour-coded MAC tape could be produced to aid drug dosing. The authors recommended a habitus modified use of Broselow tape. Broselow tape performed better in younger children, MAC better in older children. **Weaknesses:** No data presented. ★★★ Risk of bias: high
So 2009 [34]	1011	USA	P	0 to 10yrs	Broselow tape, Leffler, Theron	Yes	**Main Results:** Broselow tape most accurate weight estimation method with best accuracy in normal BMI children. Study population BMI 17.8 kg/m^2. A new formula developed, but not tested. **Weaknesses:** Poor statistical analysis and interpretation of data (bias mistaken for accuracy). Broselow tape not actually used and Broselow tape version not reported. ★★★★ Risk of bias: low
Stewart 2009 [35]	475	Australia	P	0 to 10yrs	Broselow tape	Yes	**Main Results:** Masters dissertation. The best performance of Broselow tape was between 10 and 25 kg but it was very inaccurate in children >25 kg. More accurate than age-based formulas, however. Good statistics. **Weaknesses:** Broselow tape version not reported. ★★★★★★★★ Risk of bias: low
Yamamoto 2009 [36]	542	Hawaii	P	–	Broselow tape, novel habitus-specific equation	No	**Main Results:** Adding body habitus data to length increases accuracy of weight estimation. Overweight children were preferentially selected in this study **Weaknesses:** Atypical statistical analysis – limited conclusions can be drawn from data. Data unable to be imputed. Broselow tape version not reported and Broselow tape not actually used. ★★★ Risk of bias: high

Table 1 (*Continued*)

(a)

Author, date	Study size (N)	Country	Design	Patient ages	Estimation techniques evaluated	Included in meta-analysis	Main results, Newcastle-Ottawa rating, Risk of Bias assessment
Casey 2010 [37]	1426	Australia	P	0 to 14yrs	EPLS, BG, Broselow tape	Yes	**Main Results:** BG was accurate in children 1–4yrs of age. The EPLS formula was the least accurate of all methods. The authors recommend that the accuracy and ease of use of the Broselow tape mandates its use in the ED and pre-hospital. The authors' suggestions that the BG was more "accurate" than the Broselow tape is misleading: the Broselow tape actually significantly outperformed the BG formula. BMI 17.9 kg/m^2 in study population. **Weaknesses:** Limited and incomplete statistics. Broselow tape not actually used and Broselow tape version not reported. ★★★★ Risk of bias: low
Cattermole 2010 [38]	1370	Hong Kong	P	0 to 11yrs	EPLS, Broselow tape, MAC formula	Yes	**Main Results:** MAC formula outperformed Broselow tape in children >5yrs. MAC was actually only more accurate in children 9 to 11yrs who were too tall for the Broselow tape. Poor accuracy of all systems demonstrated. **Weaknesses:** Broselow tape not actually used and 1998 edition values used. The reference for desired accuracy targets not based on evidence. ★★★★★★★ Risk of bias: low
Lulic 2010 [39]	209	Croatia	P	0 to 14yrs	EPLS, Argall, BG, Luscombe, Broselow tape	Yes	**Main Results:** Abstract. The Broselow tape was more accurate than age-based formulas but tended to slightly underestimate weight. None of the systems estimated weight with an acceptable degree of accuracy. **Weaknesses:** Broselow tape version not reported. Limited statistical analysis and data presentation. ★★★★★★ Risk of bias: low
Bourdeau 2011 [40]	243	Canada	R	1 to 10yrs	Broselow tape	Yes	**Main Results:** First Nations children selected with high incidence of obesity in this population. The Broselow tape was not accurate in First Nations children and may need to increase weight estimates by 12%. The authors comment that although ideal method of dosing still unknown, it is best practice to eliminate any possible contributing inaccuracies. Broselow tape did not perform equally well in all ethnic/population groups. **Weaknesses:** Incorrect use of some statistics. Broselow tape not actually used and Broselow tape version not reported. ★★★★★★ Risk of bias: low
Geduld 2011 [41]	2832	South Africa (Cat II)	P	0 to 10yrs	EPLS, Luscombe, BG, Broselow tape	Yes	**Main Results:** Data from a poor community in South Africa. The Broselow tape and EPLS formula were most accurate in this population. The accuracy of EPLS formula was the best ever reported while that of the Broselow tape was on par with other reports. Only 4% of children excluded as too tall for the tape. The authors suggest that careful titration of drugs and use of clinical judgment most important in using medications safely and whether the differences in accuracy of any weight estimation system are likely to affect outcomes. **Weaknesses:** Broselow tape not actually used and Broselow tape version not reported. ★★★★★★★ Risk of bias: low
Knight 2011 [42]	657	USA	R	–	Broselow tape	Yes	**Main Results:** Broselow tape 2007 B used. The Broselow tape performed poorly, potentially leading to under-resuscitation in all weight categories, especially in younger children. Drug doses correct in only about 50% of cases. Consensus opinion required whether to use ideal body weight or total body weight during resuscitation. There was a high incidence of obesity in the study population. **Weaknesses:** Broselow tape not actually used. Limited statistical analysis with no direct assessment of accuracy of weight estimation. ★★★★ Risk of bias: low

Table 1 (*Continued*)

(a)

Author, date	Study size (N)	Country	Design	Patient ages	Estimation techniques evaluated	Included in meta-analysis	Main results, Newcastle-Ottawa rating, Risk of Bias assessment
Rosenberg 2011 [43]	372	USA	P	0 to 14yrs	Healthcare provider guesses, Broselow tape	Yes	**Main Results:** Broselow tape 2007 B used. The Broselow tape was better than guesses by doctors, but not in obese children. Doctors assessed habitus poorly. Ideal body weight. Nearly 35% of study population was obese or overweight. Mean BMI 17.4 kg/m^2. It was a reflection of how poorly the Broselow tape performed in obese children that doctor estimates were better; the overall findings of Broselow tape accuracy similar to other studies. **Weaknesses:** Incomplete data presentation. Incorrect usage of some statistics. ★★★★★★ Risk of bias: low
Abdel-Rahman 2012 [44]	1938	USA	R	0 to 16yrs	EPLS, ARC, Argall, BG, Broselow tape, Cattermole, Leffler, Luscombe, Nelson, Shann, TJ, TK, Mercy method	Yes	**Main Results:** The Mercy method performed significantly better than the age-based, length-based and MAC-based systems. This was the original development and validation study of the Mercy method. Good mix of overweight children, but few underweight kids in sample. The Broselow tape was unable to be used in more than one third of the sample (too tall for the tape). **Weaknesses:** Broselow tape not actually used and Broselow tape version not reported. ★★★★★★ Risk of bias: low
Heyming 2012 [45]	491	USA	P	IQR 10 to 49mo	Broselow tape EMS vs. Broselow tape ED	No	**Main Results:** The Broselow tape was an accurate tool to estimate weight. Agreement between pre-hospital personnel and ED personnel for Broselow tape weight was only 70.1%. **Weaknesses:** Most of study population was <4yrs. Incorrect use of statistics for assessing accuracy of weight estimation. Broselow tape version not reported. ★★★★ Risk of bias: high
Milne 2012 [4]	6361	Canada	R	IQR 1.6–7.7yrs	Broselow tape	Yes	**Main Results:** Broselow tape edition 2002A used. The Broselow tape was effective, although significantly underestimated weight. Similar findings in urban and rural children. The authors suggested that an error range of 30% is reasonable and that, ideally, technology should be developed to estimate weight and aid with drug dosing. **Weaknesses:** Broselow tape not actually used – data taken from anthropometric measurements. ★★★★★★ Risk of bias: low
Park 2012 [46]	124095	Korea	R	0 to 14yrs	EPLS, Shann, Leffler, Nelson, BG, Broselow tape, Park	Yes	**Main Results:** All methods underestimated weight and performed poorly, possibly because of a secular trend towards increasing obesity. The Park formula performed best and was accurate in children <1yr of age. **Weaknesses:** Limited statistical analysis and poor interpretation of bias vs. accuracy. Broselow tape was not actually used and Broselow tape version not reported. ★★★★ Risk of bias: low
Sinha 2012 [47]	118	USA	P	0 to 14yrs	Broselow tape	Yes	**Main Results:** Broselow tape weight was compared to stretcher weight. It was possible to weigh children during trauma resuscitation, but nearly 40% of children in the study could not be weighed – more often the sicker patients (the authors did not fully explain this). There was a poor performance by Broselow tape and 18% of children were too tall for the tape. **Weaknesses:** There was no validation of the accuracy of weight measured during resuscitation. Incomplete statistics and data presentation. Broselow tape version not reported. ★★★★★★ Risk of bias: low
Trakulsrichai 2012 [48]	595	Thailand (Cat II)	P	0 to 12yrs	Broselow tape, parental estimates, growth charts	Yes	**Main Results:** Family member estimation was most accurate and the Broselow tape the most accurate of other weight estimation methods. There was equal underestimation and overestimation by Broselow tape while family estimates tended to overestimate. **Weaknesses:** Incomplete statistics. Broselow tape version not reported. Risk of bias: low

Table 1 (*Continued*)

(a)

Author, date	Study size (N)	Country	Design	Patient ages	Estimation techniques evaluated	Included in meta-analysis	Main results, Newcastle-Ottawa rating, Risk of Bias assessment
Wozniak 2012 [49]	777	Botswana (Cat III)	P	18mo to 12yrs	EPLS, Luscombe, Theron, Cattermole, Broselow tape	Yes	**Main Results:** Masters dissertation. The prediction models incorporating MAC and either tibia or ulna length performed extremely well. Age-based formulas were very inaccurate. Weight was markedly overestimated by the Broselow tape and by age-based formulas in this population with a high prevalence of HIV. **Weaknesses:** Incomplete statistics. ★★★★★★★ Risk of bias: low
Abdel-Rahman 2013 [50]	624	USA	P	0 to 16yrs	Mercy method 2D, 3D, Broselow tape	Yes	**Main Results:** Multicentre study (3) in the USA. Both 2D and 3D Mercy methods outperformed the Broselow tape although the 3D tape was significantly less accurate than the 2D method. The study population had a mean BMI of 17.3 kg/m^2. Children much older than most weight-estimation studies were included. One third of children were too tall for the tape. **Weaknesses:** Broselow tape not actually used and Broselow tape version not reported. ★★★★★★★ Risk of bias: low
Abdel-Rahman 2013 [51]	976	USA	P	0 to 16yrs	Mercy method, EPLS, Broselow tape, Luscombe, Nelson	Yes	**Main Results:** The Mercy method was accurate across a wider age range than other methods. The study population mean BMI was 17.6 kg/m^2 with very few underweight children. Inter-rater assessment was generally reasonable, but two of the raters were significantly inferior to the others. **Weaknesses:** Broselow tape not actually used and Broselow tape version not reported. ★★★★★★★ Risk of bias: low
Akabarian 2013 [52]	403	Iran (Cat II)	P	0 to 14yrs	Broselow tape, parental estimates	Yes	**Main Results:** Article in Arabic. The Broselow tape was less accurate than parental estimates, but recommended for use by the authors. Very good performance of Broselow tape compared to previous studies. **Weaknesses:** Exclusion criterion of weight >35 kg skewed the assessment of Broselow tape accuracy. Limited statistical analysis and data presentation. Broselow tape version not reported. ★★★ Risk of bias: unknown
Cattermole 2013 [53]	171	Hong Kong	P	7 to 11yrs	Broselow tape	No	**Main Results:** Selective evaluation of children who were too tall for Broselow tape. Children "too tall for the tape" do not have full adult weight – this assumption would lead to an average 30% overestimation of weight in this study. About 40% of children were too tall for the tape. Interesting data, which highlights a flaw in the Broselow tape methodology. Broselow tape was not practically useful over the age of 10. An overestimation of 30% would be unacceptable but there was no mention of desirable target. **Weaknesses:** Relatively small sample size. ★★★★★★ Risk of bias: high
Graves 2013 [54]	37114	USA	R	0 to 14yrs	EPLS, new APLS, BG, Luscombe, Broselow tape	Yes	**Main Results:** The Broselow tape 2007 B was used. The new APLS formula was best for infants, BG best for older children. The Broselow tape placed up to 60% of children in the wrong zone. Despite criticisms of Broselow tape, it outperformed the formulas in every analysis. This study had some of the poorest performances of aged-based formulas in any study. A calculation of the study population BMI was 18–22 kg/m^2. **Weaknesses:** Incomplete and poorly interpreted statistics. Broselow tape not actually used. ★★★★★★ Risk of bias: low
House 2013 [55]	967	Kenya (Cat III)	P	0 to 14yrs	Broselow tape, EPLS, Nelson	Yes	**Main Results:** The Broselow tape 2007 B was used. The Broselow tape performed better than the formulas and a measure of habitus assessment (e.g. MAC) was suggested to improve the Broselow tape. Underestimation of weight predominated. **Weaknesses:** Incomplete statistical analysis and data presentation. Poor interpretation of results from previous studies; flawed outcome measures used (indicators of bias only). ★★★★ Risk of bias: low

Table 1 (*Continued*)

(a)

Author, date	Study size (N)	Country	Design	Patient ages	Estimation techniques evaluated	Included in meta-analysis	Main results, Newcastle-Ottawa rating, Risk of Bias assessment
Loo 2013 [56]	875	Singapore	P	1 to 10yrs	EPLS, Luscombe, Broselow tape	Yes	**Main Results:** The Broselow tape 2007 B was used. The Broselow tape was more accurate than formulas. EPLS underestimated, Luscombe overestimated weight, but EPLS performed better overall than Luscombe. Mean BMI of study population 15.9 kg/m^2. Good statistics. **Weaknesses:** Broselow tape not actually used. ★★★★★★★ Risk of bias: low
Suh 2013 [57]	105072	Korea	R	–	Broselow tape 2005A, 2007B, 2011A	Yes	**Main Results:** Abstract. Large retrospective database sample. The Broselow tape 2011A was more accurate than older versions. **Weaknesses:** Statistical analysis incomplete. Broselow tape not actually used. ★★★★★★ Risk of bias: low
Wells 2013 [58]	453	South Africa	P	0 to 12yrs	Broselow tape, PAWPER tape	Yes	**Main Results:** The Broselow tape 2007 B was used. Multi-centre study of habitus-modified length-based weight estimation. Population with mixed under- and overweight. The PAWPER tape performed better than Broselow tape in every category analysed and was better than any previously published system. **Weaknesses:** Assessment of body habitus based on visual assessment. ★★★★★★★★ Risk of bias: low
Young 2013 [59]	207	USA	P	1 to 9yrs	EPLS, Broselow tape, parental estimates, Luscombe, finger counting	Yes	**Main Results:** Finger counting system as accurate as Broselow tape and more accurate than other formulas. Conceptually a simple system. To increase the accuracy weight-estimation systems may cause increased complexity and stress during resuscitations. The finger counting method is equivalent to formula Wt = 2.5 × age(yrs) + 7.5. Median BMI of study population was 17.2 kg/m^2. **Weaknesses:** Limitations in statistical analysis. Broselow tape not actually used and version not reported. ★★★★★★★ Risk of bias: low
Allison 2014 [60]	2102	Australia	R	0 to 5yrs	EPLS, BG, Luscombe, Argall, Nelson, Broselow tape, Sandell tape	Yes	**Main Results:** The Broselow tape 2007 B was used. Aboriginal and Torres Strait Island children included (low weight-for-length). The Broselow tape was most accurate in this study. Only published study of Sandell tape. **Weaknesses:** Broselow tape not actually used. Limited statistics and data presentation. Narrow age range evaluated. Very fat and very thin children excluded from study. ★★★★ Risk of bias: unknown
Batmanabane 2014 [8]	374	India (Cat III)	P	0 to 16yrs	EPLS, ARC, Argall, BG, Broselow tape, Cattermole, Leffler, Luscombe, Nelson, Shann, TJ, TK, Mercy method	Yes	**Main Results:** The Mercy method performed well in Indian children, similar to that shown in Western populations. Poor performance of the Broselow tape. Good statistics. Overestimation of weight by all methods except Mercy. **Weaknesses:** Broselow tape not actually used and Broselow tape version not reported. ★★★★★★★ Risk of bias: low
Chiengkriwate 2014 [61]	3869	Thailand (Cat II)	R	0 to 15yrs	Broselow tape	Yes	**Main Results:** The Broselow tape 2007 edition A used. The Broselow tape underestimated weight in Thai children, more so in older children, similar to findings in Western populations: Broselow tape performance consistently at a PW10 of just below 60%. **Weaknesses:** Broselow tape not actually used. ★★★★★★ Risk of bias: low
Dicko 2014 [7]	473	Mali (Cat III)	P	0 to 16yrs	Mercy, EPLS, ARC, Broselow tape, Nelson	Yes	**Main Results:** Mercy method performed extremely well in this population in Mali, similar to its performance elsewhere in the world. Other methods, including the Broselow tape, overestimated weight. BMI of study population 15.6 kg/m^2, with 22% underweight and 1.7% overweight or obese. Good inter-rater reliability. **Weaknesses:** Some limitations in statistics. Broselow tape not actually used and version not reported. ★★★★★★ Risk of bias: low

Table 1 (*Continued*)

(a)

Author, date	Study size (N)	Country	Design	Patient ages	Estimation techniques evaluated	Included in meta-analysis	Main results, Newcastle-Ottawa rating, Risk of Bias assessment
Asskaryar 2015 [62]	1185	India (Cat III)	P	1mo to 12yrs	Broselow tape	Yes	**Main Results:** The Broselow tape 2007 edition B was used. The Broselow tape significantly overestimated weight in Indian children. The authors claimed that an 8% modification of the tape improved its accuracy, although the improved performance was not substantially better than the original and was still below acceptable performance. Recalibrating bias alone is not enough when precision is low. **Weaknesses:** Limited data presentation and statistical analysis. ★★★★★ Risk of bias: low
Britnell 2015 [63]	376	New Zealand	P	5 to 10yrs	EPLS, Shann, Theron, Broselow tape	Yes	**Main Results:** The Broselow tape 2011 edition A was used. The Broselow tape was more accurate than age-based formulas in children <143 cm. Current acceptance of formulas needs to change. Large differences in accuracy of weight estimation shown in different ethnic groups. There was a large proportion of Pacific Island children in sample. Broselow tape could not be used in one fifth of children, but best ever performance of the Broselow tape in a study. Habitus was assessed but data not presented. **Weaknesses:** Limited and incomplete statistics. ★★★★★ Risk of bias: low
Chavez 2015 [64]	324	USA	P	1mo to 12yrs	Broselow tape, PAWPER, APLS, MAC	Yes	**Main Results:** Age-based formulas and the MAC formulas performed badly. The Broselow tape performed better and the PAWPER was most accurate overall, although not as accurate as in previous studies. Underestimation of obesity (and habitus score) caused underestimation of weight. High level of obesity in study population. **Weaknesses:** Broselow tape version not reported. Limited and incomplete statistics. PW5 mistaken for PW10 by authors – confusing for readers. ★★★★★ Risk of bias: low
Georgoulas 2015 [65]	300	South Africa (Cat III)	P	1mo to 12yrs	Broselow tape, PAWPER, Wozniak, Mercy method	Yes	**Main Results:** The Broselow tape 2011 edition A was used. Poor population with high proportion of underweight children. First comparative study of these methodologies. The PAWPER tape performed best, but good performances from Wozniak and Mercy methods. The Broselow tape was the worst performer. Relatively weaker performance by all methods in infants, but Wozniak especially was very weak. **Weaknesses:** Assessment of body habitus done by single researcher. ★★★★★★★★ Risk of bias: low
Khouli 2015 [66]	815	Mexico (Cat II)	P	0 to 12yrs	Broselow tape	Yes	**Main Results:** The Broselow tape was not accurate in Mexican population. Reasonable statistics. Both under- and overestimation found to be a problem. Nearly half of children had at least one colour zone error. **Weaknesses:** Broselow tape not actually used and Broselow tape version not reported. Poor interpretation of findings. ★★★★★ Risk of bias: low
O'Leary 2015 [67]	199	Australia	P	0 to 14yrs	PAWPER, APLS, Luscombe, BG, Broselow tape, Mercy method	Yes	**Main Results:** Age-based formulas performed badly. The Broselow tape and Mercy method performed significantly better and the PAWPER was most accurate overall. All systems performed worst in infants. There was a reasonably good performance of the Broselow tape, possibly because of newer edition used. **Weaknesses:** Broselow tape version not reported. ★★★★★★ Risk of bias: low
Young 2015 [68]	207	Philippines (Cat III)	P	1 to 9yrs	EPLS, APLS, Luscombe, BG, finger counting, Broselow tape	No	**Main Results:** The Broselow tape 2011 edition A was used. The Broselow tape performed best in this population, and the updated APLS formula performed worst. **Weaknesses:** Broselow tape not actually used. Very limited and incomplete statistical analysis. Does not allow for comparison to other studies or for inclusion into meta-analysis. ★★★★ Risk of bias: low

Table 1 (*Continued*)

(a)

Author, date	Study size (N)	Country	Design	Patient ages	Estimation techniques evaluated	Included in meta-analysis	Main results, Newcastle-Ottawa rating, Risk of Bias assessment
Abdel-Rahman 2015 [69]	400**	USA	P	2mo to 16yrs	EPLS, Luscombe, DWEM, healthcare provider guesses, Broselow tape, 2DMT, 3DMT	Yes	**Main Results**: "Real world" performance of weight estimation less accurate than when performed by experts. Calculation errors were common with age-based formulas. Weight and habitus were often underestimated by visual inspection. Usage errors with Broselow tape and Mercy tapes were common. **Weaknesses:** Not all modern weight estimation systems evaluated. ★★★★★★ Risk of bias: low
Talib 2015 [70]	318	USA	P	0 to 18yrs	APLS, Broselow tape, MAC formula, Mercy method	No	**Main Results**: The Mercy method performed better than Broselow tape and age-based formulas in children with Down syndrome. No system performed well in this population. **Weaknesses:** Incomplete statistics obscure the poor performance of all methods, including the Mercy method. ★★★ Risk of bias: high
AlHarbi 2016 [71]	3537	Saudi Arabia	P	1mo to 12yrs	Broselow tape 2007 B and Broselow tape 2011A	No	**Main Results**: The Broselow tape 2011A performed better than Broselow tape 2007 B in this population. The authors suggested that the tapes were accurate. The method of statistical analysis does not support the conclusions drawn in this paper. **Weaknesses:** It was unclear if the Broselow tape was used or if Broselow tape weights were derived from length measurements. Limited, incomplete and inappropriate statistical analysis. ★★★ Risk of bias: high
Aliyu 2016 [72]	300	Nigeria (Cat II)	P	0 to 5yrs	Broselow tape, APLS	Yes	**Main Results**: The Broselow tape and APLS formulas performed well in this population. Contrary to the authors conclusions, although the Broselow tape and APLS formula performed similarly, they were both inaccurate. **Weaknesses:** Only young children included. Not clear if Broselow tape used, or if derived from length measurements. Broselow tape edition not reported. Incomplete statistical analysis. ★★★★★ Risk of bias: low
Bowen 2016 [73]	1381	Zambia (Cat II)	P	0 to 14yrs	Broselow tape, APLS, EPLS, ARC, Argall, BG, Garwood, Luscombe, Michigan, Nelson, Park, Shann, Theron, Tintinalli	Yes	**Main Results**: Broselow tape performed better than every formula in this population, BG and Michigan formulas performed worst. None of the methods were accurate and all methods overestimated weight. **Weaknesses:** Broselow tape not actually used. Broselow tape version not reported. ★★★★★★ Risk of bias: low
Clark 2016 [6]	583	Sudan (Cat III)	R	6mo to 5yrs	Broselow tape	Yes	**Main Results:** Study in South Sudan, the "hungriest place on earth" where 61% of study population was malnourished. The Broselow tape performed very poorly. There was up to two colour zone overestimation in severely malnourished children with only 26% agreement in normally nourished children: a potentially dangerous overestimation of weight in undernourished children. **Weaknesses:** Very limited statistical analysis and data presentation. Broselow tape not actually used and Broselow tape version not reported. ★★★★ Risk of bias: low
Jung 2016 [74]	906	Korea	P	0 to 17yrs	Broselow tape, novel device	Yes	**Main Results**: The novel device performed better than the Broselow tape in all outcome measures and was quicker to use. Both devices underestimated weights. Overweight and underweight children were often misclassified into wrong habitus category. Broselow tape 2011 edition A. **Weaknesses:** Device is not commercially available. Incomplete statistical analysis. ★★★★★★ Risk of bias: low
Jung 2016 [75]							Non-systematic review. **Main Results:** Weight frequently underestimated in older children. Differences in body habitus account for this error, which needs further research. **Comments**: Article in Korean. **Weaknesses:** Non-systematic review. No objective assessment of methodologies.

Table 1 (*Continued*)

(a)

Author, date	Study size (N)	Country	Design	Patient ages	Estimation techniques evaluated	Included in meta-analysis	Main results, Newcastle-Ottawa rating, Risk of Bias assessment
Lowe 2016 [76]	3018	USA	R	0 to 13yrs	Broselow tape, Handtevy tape	Yes	**Main Results**: The Broselow tape performed better than the Handtevy tape in all outcome measures. Both tapes underestimated weights. Although the authors recognised the increased accuracy when obese children were excluded, they still erroneously advocated "recalibration" of the tapes. Broselow tape 2011 edition A. **Weaknesses:** Tapes not actually used. Incomplete statistical analysis. ★★★★★ Risk of bias: low
Mishra 2016 [77]	603	India (Cat III)	P	0 to 10yrs	Broselow tape	No	**Main Results**: Broselow tape performed best in smallest children. Broselow tape 2007 edition B. **Weaknesses:** Limited and incomplete statistical analysis. ★★★ Risk of bias: high
Nosaka 2016 [78]	237	Japan	P	0 to 10yrs	EPLS, Park, Broselow tape, parental estimates	Yes	**Main Results**: Parental estimates were most accurate, then Broselow tape then APLS formula. Mothers' weights, as well as Broselow tape and EPLS estimations were extremely accurate. The BMI for this population suggests that very few obese children were included and the sample was skewed towards very young children. Broselow tape 2007 edition B. **Weaknesses:** Limited and incomplete statistical analysis. ★★★★★ Risk of bias: low
Ralston 2016 [79]	453990	Multicentre International (Cat II)	R	6mo to 5yrs	Broselow tape, novel MAC-based formulas	Yes	**Main Results**: A model incorporating height and MAC was the most accurate. Broselow tape 2011 edition A less accurate than 2007 edition B in this population. This was a very large multinational database study. **Weaknesses:** Broselow tape not actually used. Only very young children <5years of age were included. ★★★★★ Risk of bias: low
Sahar 2016 [80]	1163	Malaysia (Cat II)	P	0 to 12yrs	Broselow tape	Yes	**Main Results**: Broselow tape underestimated weight in small children and overestimated weight in older children. It was not accurate overall. As with studies elsewhere there was a large variation in accuracy. **Weaknesses:** Broselow tape version not reported. Limited and incomplete statistical analysis. ★★★★★ Risk of bias: low
Tanner 2016 [81]	178	USA	P	2 to 18yrs	Broselow tape	No	**Main Results**: The Broselow tape was very inaccurate in overweight and obese children, who made up more than half the sample. Habitus assessment was poor by parents and nurses, but better by the principal investigator. Using a corrective formula based on waist circumference to adjust the Broselow tape weight improved accuracy in obese children. **Weaknesses:** Broselow tape version not reported. Limited, unclear and incomplete staually used. Only very young children <5years of age were included. ★★★★★ Risk of bias: low
Sahar 2016 [80]	1163	Malaysia (Cat II)	P	0 to 12yrs	Broselow tape	Yes	**Main Results**: Broselow tape underestimated weight in small children and overestimated weight in older children. It was not accurate overall. As with studies elsewhere there was a large variation in accuracy. **Weaknesses:** Broselow tape version not reported. Limited and incomplete statistical analysis. ★★★★★ Risk of bias: low
Tanner 2016 [81]	178	USA	P	2 to 18yrs	Broselow tape	No	**Main Results**: The Broselow tape was very inaccurate in overweight and obese children, who made up more than half the sample. Habitus assessment was poor by parents and nurses, but better by the principal investigator. Using a corrective formula based on waist circumference to adjust the Broselow tape weight improved accuracy in obese children. **Weaknesses:** Broselow tape version not reported. Limited, unclear and incomplete statistical analysis. Only included children who fell within the length limitations of the tape. ★★★ Risk of bias: high
Young 2016 [82]							Systematic review. **Main Results:** Parental estimates are the most accurate technique, followed by length- and habitus-based methods. Broselow tape more accurate than age formulas. **Weaknesses:** Limited evaluation of data in this review.

Table 1 (*Continued*)

(a)

Author, date	Study size (N)	Country	Design	Patient ages	Estimation techniques evaluated	Included in meta-analysis	Main results, Newcastle-Ottawa rating, Risk of Bias assessment
Waseem 2017 [12]	538	USA	R	0 to 8yrs	Broselow tape	No	**Main Results**: Broselow tape underestimated weight in older children. It was not accurate in nearly half the population. Broselow tape 2011 edition A. Many underweight and obese children included which resulted in poor performance. **Weaknesses:** Broselow tape not actually used. Limited and incomplete statistical analysis. Only included children who fell within the length limitations of the tape. ★★★ Risk of bias: high
Wells 2017 [83]	328	South Africa	P	0 to 16yrs	Broselow tape, PAWPER, Wozniak, Mercy method	Yes	**Main Results:** This was a population selected for older children and for children with deviations from "average" weight-for-length. The Broselow tape performed poorly in this study, the Wozniak method and the Mercy method showed intermediate accuracy and the PAWPER was most accurate overall. The extended PAWPER tape worked well in this population. **Weaknesses:** The Mercy method was used in a simulated resuscitation setting (supine children), which may have affected its accuracy. ★★★★★★★ Risk of bias: low
Wells 2017 [84]	1085	South Africa	P	0 to 16yrs	Broselow tape	No	**Main Results:** Habitus-modified models greatly improved the accuracy of the Broselow tape. **Weaknesses:** The visual assessment of habitus was performed by moderately experienced users. ★★★★★★★ Risk of bias: high
Whitfield 2017 [85]							As for Wozniak 2012.

(b)

Author, date	Size (N)	Design	Endpoints	Risk of Bias	Major findings; comments; major limitations
Luten 1992 (USA) [86]	213	Prospective derivation and validation of length-based ETT size estimation	Endotracheal tube size	Low	**Main Results:** The Broselow tape predicted correct uncuffed endotracheal tube size in 77% of children and 99% within 0.5 mm of correct size. **Comments**: One of the only studies to objectively evaluate the accuracy of equipment size prediction by the Broselow tape. Other equipment sizes determined by modified Delphi method. **Weaknesses:** The relevance of uncuffed endotracheal tube sizes is far less now than 25 years ago. ★★★★★★★
Lanoix 1999 (USA) [87]	N/A	Clinical communication		–	**Main Results:** Colour-coding organisation of resuscitation equipment facilitates emergent medical management. **Comments**: Theoretical concept proposition. **Weaknesses:** No evidence presented.
Hohenhaus 2001 [88]	N/A	Clinical report		–	**Main Results:** Broselow tape recommended for use, but many errors when used by healthcare providers with insufficient training. **Comments**: Broselow tape helpful only when used correctly. **Weaknesses:** No quantitative data presented.
Vilke 2001 [89]	20	Prospective simulation study evaluating accuracy of adrenaline dose by paramedics	Adrenaline dose	High	**Main Results:** The authors claim that 95% of estimates were within acceptable error range, but there was an unrealistic target range, with no supporting evidence basis. Tenfold errors in drug doses occurred in 10% of cases. The Broselow tape was more accurate than guesses. **Weaknesses:** Very poor statistics. Limited data analysis made findings unreliable and uninterpretable. ★★★★
Luten 2002 (USA) [90]	N/A	Clinical communication/editorial			**Main Results:** Drug dose calculations must be eliminated and a global strategy to limit extraneous thinking during emergency management is required. **Comments**: A review of cognitive stressors experienced during paediatric resuscitation. **Weaknesses:** No data presented.
Slishman 2002 (USA) [91]	7	Prospective simulation study comparing standard Broselow tape techniques and a new device, the "per-kilo doser"	Time to drug withdrawal	Unknown	**Main Results:** Drug administration faster with new "per-kilo doser" device, by about 30 s per drug. **Comments**: 7 nurses each preparing 4 medications using standard Broselow tape techniques and 4 using a new device, the "per-kilo doser" Broselow tape colour zones were used for both systems. **Weaknesses:** Small study. No standardised control group. ★★★★★★
Hohenhaus 2003 (USA) [92]	N/A	Abstract		–	**Main Results:** The Broselow tape requires training for proper use. **Comments**: Reiteration of the fact that the Broselow tape is not as easy to use as might be expected. **Weaknesses:** No data presented.

Table 1 (*Continued*)

(b)

Author, date	Size (N)	Design	Endpoints	Risk of Bias	Major findings; comments; major limitations
Shah 2003 (USA) [93]	28	Prospective simulation crossover study with 28 doctors comparing traditional dosing system and colour-coded system.	Deviation from recommended dose range and deviation from recommended equipment sizes	Low	**Main Results:** Colour-coding significantly reduced the number of errors. Participants preferred the Broselow tape system to the standard system. **Comments**: Supplementary information is essential to the success of the Broselow tape. **Weaknesses:** Colour-coded drug sheets used in addition to the Broselow tape. ★★★★★★★
deBoer 2004 (USA) [94]	N/A	Clinical communication.			**Main Results:** Colour-coding can be used for other applications e.g. spinal immobilisation appropriate for size. **Comments**: An integrated approach to all aspects of the management of the critically ill or injured child is essential. **Weaknesses:** No data presented.
Hohenhaus 2004 (USA) [9]	N/A	Clinical communication.		–	**Main Results:** The Broselow tape may cause significant weight estimation errors if used incorrectly. The Broselow tape is intended more for equipment size determination than weight estimation. **Comments**: Some errors and weaknesses of the 1998 edition highlighted. **Weaknesses:** Limited quantitative data presented.
Moore 2004 (USA) [95]	144	Prospective simulation study assessing the accuracy of Broselow tape use by paramedics	Dose of resuscitation drugs	Unknown	**Main Results:** Despite using Broselow tape, only 56% correct medication doses delivered. Despite improving weight estimation, the Broselow tape did not decrease medication errors. **Weaknesses:** Limited data analysis. ★★
Agarwal 2005 (USA) [96]	21	Prospective crossover simulation trial with 21 paediatric-trained healthcare providers comparing standard resuscitation cart with colour-coded Broselow cart.	Time to and accuracy of selection of appropriate medical equipment along with participant satisfaction	Low	**Main Results:** The Broselow cart was preferred by two thirds of participants. It was slightly faster and slightly more accurate than a standard cart. **Comments**: Although the accuracy of the Broselow cart was reported to be more accurate, review of the actual data shows small and inconsistent effect size. **Weaknesses:** Small sample. Only equipment was evaluated. The importance of the sizes of much of the equipment evaluated is debatable. ★★★★★★★
deBoer 2005 (USA) [1]	N/A	Clinical communication.			**Main Results:** A patient can be "bumped up" a zone if they are obese, except for medications that should be dosed according to lean or ideal body weight. Colour-coded reference materials should be used for drug dose guidelines. **Comments**: Repetition of previous articles with no evidence-based material. **Weaknesses:** No data presented to support content. Over-reliance on reference to opinion-based publications of the developers.
deBoer 2005 (USA) [97]	N/A	Clinical communication.			**Main Results:** "In most cases of emergency resuscitation or stabilisation, the medications administered to children are the same as those for adults. The dosages, however, are dramatically different" "The simplest [solution] is a book, in print or in an electronic format [. . .] with colour-coded pages. All doses are pre-calculated not only in milligrams (mg), but also in terms of the millilitres (ml) of a standard concentration". **Comments**: The Broselow tape should not be used as a stand-alone resuscitation aid. **Weaknesses:** No original data presented.
Hohenhaus 2005 (USA) [98]	N/A	Editorial/clinical communication.		–	**Main Results:** Broselow tape colour zone is a "vital sign". Delay at any point in emergency care should be considered an error. Administrators should ensure that clinicians have the tools they require to deliver paediatric-specific care. **Comments**: Further emphasis of the necessity avoid calculations and to have the right equipment available. **Weaknesses:** Opinion piece.
Frush 2006 (USA) [99]	89	Randomised control trial of Broselow tape use with and without training in its use.	Dosing deviation from accepted dose range and time to determine the dose for each medication prescribed	Low	**Main Results:** Slight increase in speed of dose prescription, with increased accuracy. New clinical tools (such as Broselow tape) may lead to new or additional errors. **Comments**: Broselow tape only used correctly in 3–5% of untrained individuals! Doses evaluated in milligram format. **Weaknesses:** No evaluation of administration-ready time i.e. time to calculate dilutions and volume to administer. ★★★★★★★
Kaji 2006 (USA) [100]	141	Observational study of Emergency Medical Services personnel's errors before and after introduction of Broselow tape into protocol	Accuracy of dose of adrenaline during cardiac arrest	High	**Main Results:** Broselow tape increased the likelihood of correct and semi-accurate dose calculation with a small to medium effect size. **Comments**: Drug dose delivery still only correct in 50% of cases with Broselow tape **Weaknesses:** Small numbers. ★★★★

Table 1 (*Continued*)

(b)

Author, date	Size (N)	Design	Endpoints	Risk of Bias	Major findings; comments; major limitations
Frush 2007 (USA) [101]	N/A	Policy statement.		–	**Main Results:** Broselow tape should be used, but "some studies have described a potential increase in medical errors when using the Broselow tape because of its design and the fact that it is often used incorrectly". **Comments**: The Broselow tape should not be used without training and without being part of a comprehensive resuscitation aid system. **Weaknesses:** No data presented.
Luten 2007 (USA) [102]	N/A	Editorial comment.			**Main Results:** No substantiation for setting acceptable accuracy of weight estimation at 10%. Broselow tape estimates ideal body weight, therefore obesity not relevant and overdose unlikely in emaciated children. **Comments**: Many comments not substantiated by evidence. **Weaknesses:** Opinion piece.
Fineberg 2008 (Australia) [103]	16	Prospective, randomised crossover trial, comparing Broselow tape with a standardised volume-weight method.	Time to medication delivery and accuracy of drug doses.	Low	**Main Results:** New method faster and with fewer errors than Broselow tape. The calculation of volumes to administer was error-prone. **Comments**: 16 participants in 3 simulations with 13 dose determinations in each arm. Volume dose calculations had to be performed. **Weaknesses:** This volume-weight method is not commercially available. ★★★★★★★
Luten 2008 (USA) [104]	N/A	Commentary in response to Fineberg 2008.			**Main Results:** Ideal body weight might be better than total body weight as a goal; "the Broselow tape is a tool that was not designed to be used without clinical judgment". **Comments**: No evidence offered to support opinion; no goal target suggested. In this letter Luten wrote, "To address these concepts, the new addition of the Broselow Tape includes the calculated volume of medication from a preferred concentration list. This requires reducing the number of medications on the tape in order to allow adequate sized font." This has not yet materialised. **Weaknesses:** This provided an important theoretical discussion, but without providing evidence.
Wells 2008 (South Africa) [105]	N/A	Letter.		–	**Comments**: The Broselow tape has not been shown to accurately predict ideal body weight. Total body weight should be used for dose calculation in emergencies. **Weaknesses:** This provided a counter view to Luten 2008, but without providing evidence.
Feleke 2009 (USA) [106]	16	Prospective sequential simulation study in which a colour-coded medication system was compared with traditional dosing references.	Time of conversion of medication orders from dose to volume and access to preparation and administration information.	Low	**Main Results:** A colour-coded system together with the Broselow tape resulted in faster time to drug administration and fewer errors in dose and administration. The Broselow tape does not provide a pre-calculated conversion from milligrams to millilitres for most medications, which may lead to errors. **Comments**: 320 medication orders were given to 16 nurses. **Weaknesses:** Colour-coded system required for benefits of Broselow tape to be realised. ★★★★★★★★
Pinchevsky 2010 (USA) [107]	4	Retrospective comparison of three dosing strategies in obese children during cardiac arrest.	Drug dose deviations	High	**Main Results:** All dosing strategies were very different and all had large errors. **Comments**: It is unclear whether this study adds anything to the field other than that there needs to be consensus on how drugs are dosed in obese children. **Weaknesses:** No clarity on how ideal body weight was determined and which dose was used as the reference dose value. Very small sample size. ★★★★

Table 1 (*Continued*)

(b)

Author, date	Size (N)	Design	Endpoints	Risk of Bias	Major findings; comments; major limitations
Kelly 2012 (Australia) [108]	N/A	Book chapter.		–	**Main Results:** Best systems to be used, in order of accuracy: parental estimates, Broselow tape, age-based formulas. **Comments**: Discussion primarily on the weight-estimation function of the Broselow tape. **Weaknesses:** No mention of acceptable targets for weight estimation or weaknesses of age formulas. Subjective assessment of studies only.
Lammers 2012 (USA) [109]	45	Prospective simulation study to detect errors in management.	Unique performance scoring protocol to determine errors during resuscitation.	Low	**Main Results:** Use of the Broselow tape alone did not eliminate errors. The Broselow tape was used incorrectly often enough to be of concern. **Comments**: A 10% error was regarded as the maximum allowable. **Weaknesses:** Limited quantitative data reported. ★★★★★
Meguerdichian 2012 (USA) [110]	N/A	Non-systematic review.		–	**Main Results:** The Broselow tape is the most consistent and reliable tool for weight estimation, but habitus-based weight adjustment for Broselow tape is logical. Parental estimates may rival the Broselow tape, but parents may be absent or uncertain, especially under stress. Underdosing might be prudent in emergencies. **Comments**: The author acknowledges that the degree of acceptable error is difficult to define. **Weaknesses:** No evidence provided of efficacy of Broselow tape in reducing medication errors. No suggestion of goal target accuracy required. Some controversial opinions about drug dosing.
Abdel-Rahman 2014 (USA) [111]	N/A	Non-systematic review.		–	**Main Results:** Some experts have suggested that weight estimation cannot be accurate. This is likely to be to the disadvantage of children as dual length-and habitus-based can achieve acceptable accuracy. **Comments**: Brief review of weight estimation systems. **Weaknesses:** No target of weight estimation accuracy suggested; Mercy method recommended for environments where no scale available – but no mention of emergency use.
Lammers 2014 (USA) [112]	142	Prospective simulation study to detect errors in management.	Performance scoring protocol used to determine errors during resuscitation.	Low	**Main Results:** Incorrect doses were found for drugs not on Broselow tape. Large errors still possible with drugs on tape. **Comments**: Simulation study with 142 participants in 62 sessions. **Weaknesses:** Limited details on quantitative data presented. ★★★★★
Moreira 2015 (USA) [10]	10	Prospective crossover simulation study comparing colour-coded syringes and conventional drug administration.	Time from initiation of medication preparation to completed administration and dose errors and critical errors.	Low	**Main Results:** Colour-coded prefilled syringes improved time to delivery of medication with a reduction in critical errors: 17% with conventional and 0% with colour-coded technique. **Comments**: of 10 doctor-and-nurse teams comparing colour-coded syringes and conventional drug administration. Within 10% were still considered an error; outside this range was a critical error. **Weaknesses:** Colour coded, prefilled syringes are not commercially available. ★★★★★★★
Stevens 2015 (USA) [113]	10	Prospective crossover simulation study comparing colour-coded syringes and conventional drug administration.	Time from initiation of medication preparation to completed administration and dose errors and critical errors.	Low	**Main Results:** Colour-coded prefilled syringes slightly decreased time to delivery of medication with a reduction in critical errors: 39% with conventional and 0% with colour-coded technique. **Comments**: Crossover simulation study of 10 paramedics comparing colour-coded syringes and conventional drug administration. **Weaknesses:** Colour coded, prefilled syringes are not commercially available. ★★★★★★★
Rappaport 2016 [114]	80	Prospective simulation study to detect errors in drug doses from Broselow tape and Handtevy system.	Drug dose errors >20%; incorrect concentration of drugs; time to administer	Low	**Main Results:** High error rates for both tapes, but especially Broselow tape (from 21% to 64% with different drugs). Tapes frequently used incorrectly. Handtevy tape allowed for faster administration of drugs. **Comments**: More information available on Handtevy tape which reduced errors. **Weaknesses:** Correct weight was defined as weight on the tapes, rather than actual weight. ★★★★★★

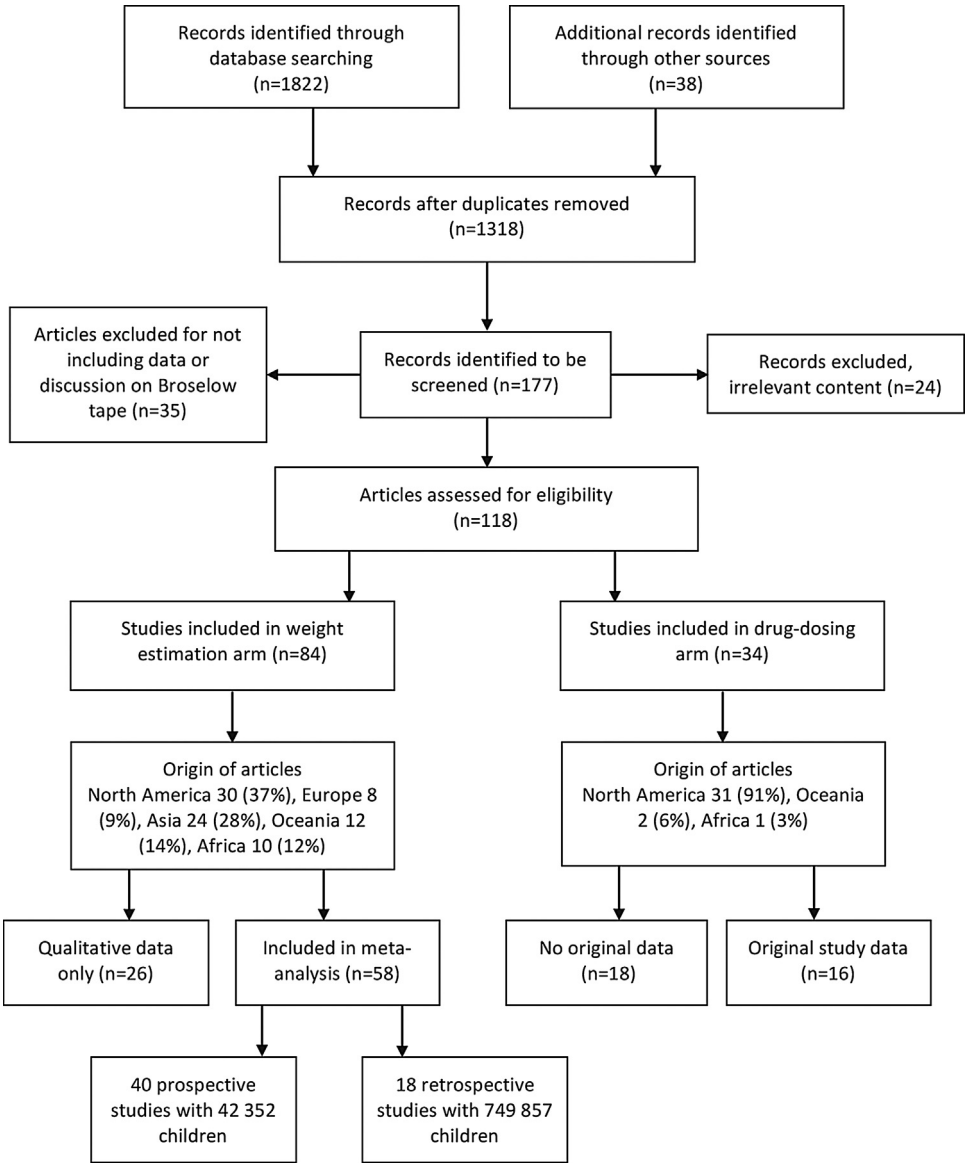


Fig. 1. Flow diagram for article identification and selection. In the "origin of articles" box, Oceania refers to Australia and New Zealand.

2.8 Software

Data were analysed and graphics generated using Microsoft excel for Mac (Microsoft. Microsoft Excel for Mac. Redmond, Washington: Microsoft, 2016), RevMan (Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014) and Prism 6 (GraphPad Prism version 6.00 for Mac, GraphPad Software, San Diego California USA, www.graphpad.com).

3 Results

3.1 Changes in the Broselow tape from 1989 to 2011

The tape has changed substantially in appearance and content over the years. The content, format and changes in the tapes from edition-to-edition are shown and described in the supplementary material (Supplementary Fig. 1 and Supplementary Table 3).

3.2 Systematic review – study selection

The details of the evaluated and included studies are shown in Fig. 1.

3.3 Meta-analysis – individual study and pooled-data findings

The bias and precision data for each study and for the pooled data analysis are shown in Fig. 2. The data on the accuracy of the Broselow tape in each included study and for the pooled data are shown in Fig. 3. The data on the performance of the tape in populations with different prevalence of underweight children are included in these figures. There was no detectable change in the performance of the Broselow tape over time or from edition to edition. The results of direct pair-wise comparisons between the Broselow tape and various other weight-estimation methods can be found in Fig. 4.

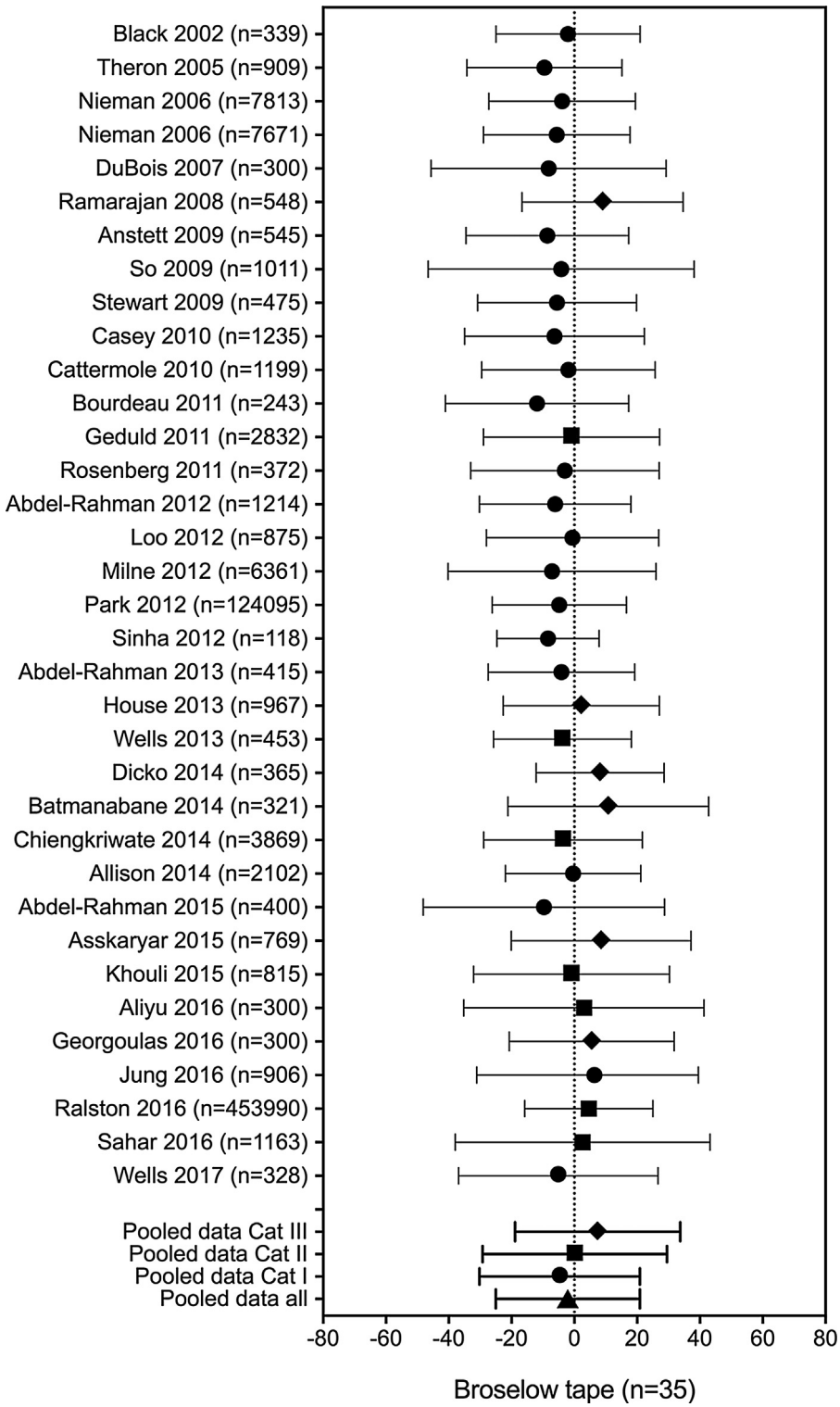


Fig. 2. Forest plot of the included studies and pooled data showing bias and 95% limits of agreement. Studies from countries with a low prevalence of underweight (Category I) are indicated by circles, studies from countries with a medium prevalence of underweight (Category II) are indicated by diamonds and studies from countries with a high prevalence of underweight (Category III) by squares.

When evaluating the performance data of the Broselow tape from individual studies, the following findings were consistently reported:

- The Broselow tape consistently was significantly *more accurate* than healthcare provider guesses (two studies) and age-based formulas (25 studies).
- Some studies showed the Broselow tape to be more accurate than parental guesses (two studies) and some showed it to be less accurate (five studies). The differences were small.
- There were mixed results when comparing the Broselow tape with the earliest dual length- and habitus-based system, the Devised Weight Estimating Method (one study favoured Broselow tape, one study favoured the Devised Weight Estimat-

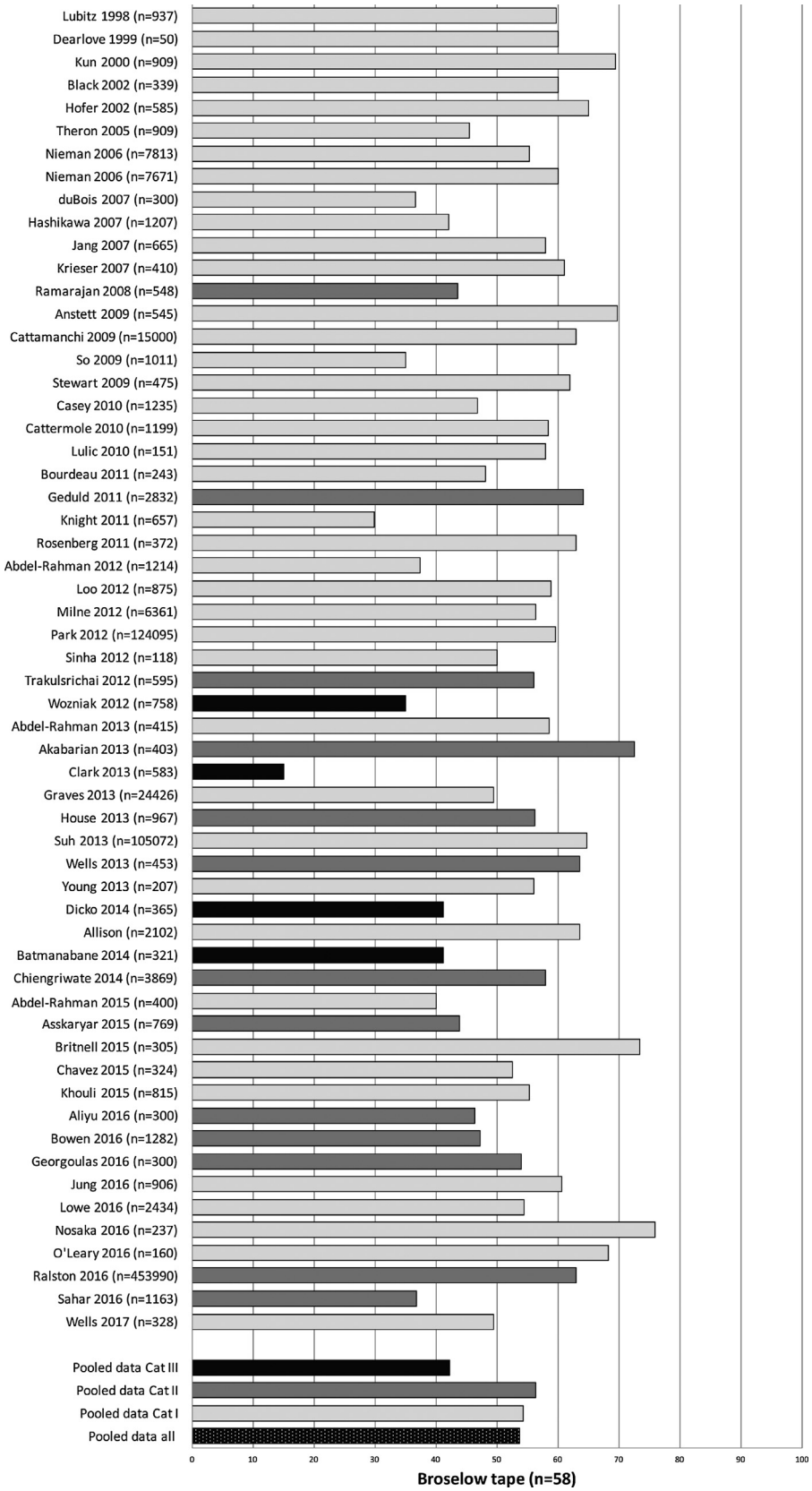


Fig. 3. Bar chart showing the accuracy data from the included studies and pooled data – percentages of weight estimations accurate to within 10% of actual weight are shown (PW10). Studies from countries with a low prevalence of underweight (Category I) are indicated by pale grey bars, studies from countries with a medium prevalence of underweight (Category II) are indicated by dark grey bars and studies from countries with a high prevalence of underweight (Category III) by black bars. There was a significant difference between the pooled accuracy of Category I and Category II studies when compared with Category III studies (Fisher exact test, $p < 0.0001$).

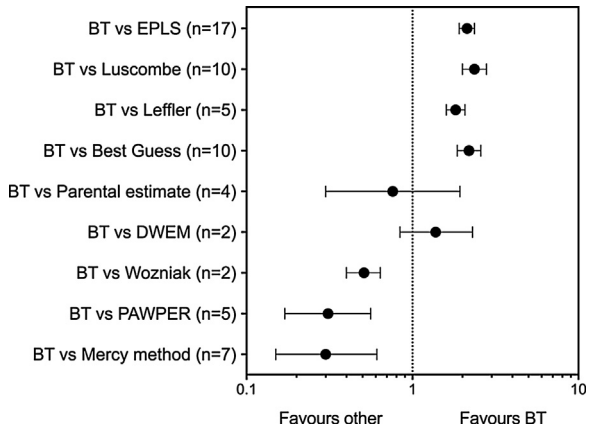


Fig. 4. Forest plot of direct comparisons between the Broselow tape and other weight estimation systems. These comparisons were made using accuracy data (percentages of weight estimations accurate to within 10% of actual weight) and inverse-variance random effects modelling. The Broselow tape was significantly *more accurate* than the four age-based formulas (for which paired data was available) with a small effect size. There was no difference in accuracy between the Broselow tape and parental estimates and the Devised Weight Estimating Method (DWEM). The Broselow tape was significantly *less accurate* than the three newer dual length-and habitus-based systems for which there was paired data.

ing Method and one reported no difference). The differences were small.
- The Broselow tape consistently performed *less accurately* than every modern dual length- and habitus-based system – the Mercy method, the PAWPER tape and the Wozniak method (14 studies). The differences had a large effect size. A description of the dual length- and habitus-based systems is contained in Supplementary Table 4.

3.4 Subgroup analysis findings

About half of the included studies (50% for bias and precision data and 42% for accuracy data) contained usable subgroup data, based on weight groupings, which can be found in Table 2. The Broselow tape was most accurate in children between 10 and 25 kg, slightly less accurate in children <10 kg and least accurate in children >25 kg.

The following main findings can be summarised from the studies that could not be incorporated into the pooled data analysis:

- In studies where children <145 cm were not excluded, the Broselow tape was frequently too short to provide a weight estimate in as many as 40% of children (seven studies) [5,41,44,47,50,53,83].
- When the Broselow tape was compared with healthcare provider guesses, the Broselow tape always performed better (two studies) [3,30].
- When the Broselow tape was compared with parental estimates, parental estimates were always more accurate (three studies) [27,28,30]; if the data from the studies included in the meta-analysis were included, parental estimates were found to be better in eight of the ten studies.
- When the Broselow tape was compared with age-based formulas, the Broselow tape was always more accurate (seven studies) [19–21,27,28,68,70].
- Three studies considered the Broselow tape to be acceptably accurate only in younger children <120 cm and <20 kg [24,33,53].
- One study showed a length-based formula modified for body habitus to be more accurate for drug dosing than the Broselow tape [36].

3.5 Drug dosing guideline findings

Of the 16 studies with original data, seven studies used time to medication delivery or medication-delivery readiness as a primary outcome:

- One study evaluated the Broselow tape before and after training and showed that training significantly improved the accuracy and speed of use of the Broselow tape [99].
- Two studies compared the Broselow tape plus colour-coded drug dosing reference material against traditional dosing references and found the Broselow system to be superior [93,106].
- Four studies compared other drug delivery aids or techniques with the Broselow tape and all four showed faster times and fewer errors than with the Broselow tape [10,91,103,113].

Twelve studies appraised the accuracy of drug dosing using the Broselow system:

- The Broselow tape was more accurate than guessing doses (one study) [89].
- The Broselow tape plus colour-coded dosing assets (e.g. drug calculation sheets) were more accurate than traditional dosing references (three studies).
- The use of the Broselow tape alone (without colour-coded dosing assets) did not eliminate errors (six studies) [10,89,95,103,112,113].
- The Handtevy tape was more accurate, with fewer errors than the Broselow tape (one study) [114].

In the 18 studies with no original data, the important, repetitive themes identified were:

- Many errors are incurred when the Broselow tape is used by untrained healthcare providers and the incorrect use of the Broselow tape may lead to errors.
- Arithmetic calculations must be avoided at every stage of the drug dose computation-preparation-administration process. The Broselow tape should not be used as a stand-alone device without using colour-coded drug assets to eliminate the need for calculations.
- The Broselow system should make use of a habitus-modified weight-estimation and drug-dosing system to improve accuracy.

4 Discussion

4.1 The impact of the changes in the Broselow tape on its function

The available evidence was not convincing that the Broselow tape's weight estimation accuracy has improved significantly over the years. This has led some authors to question whether the tape is still valid, as there is evidence that both its weight estimation and drug dosing roles are suboptimal and have not successfully been adapted to a changing population [4,5]. The most concerning discovery, however, is that the most recent edition of the tape is significantly less accurate than previous editions in low- and middle-income countries, with increased overestimation of weight, which could lead to significant medication overdosing errors [79].

4.2 The Broselow tape as a weight estimation device

The bias and precision data showed a consistently high within-population variance, which was similar to the between-population variance. This suggests that, even in the same study population, the Broselow tape was unable to estimate weight consistently in chil-

Table 2
Summary of the pooled data including subgroup analyses for the accuracy, bias and precision of the Broselow tape.

	Subgroup		MPE Mean (95%CI)	LLOA (95%CI)	ULOA (95%CI)	Number of children (number of studies)	PW10 (95%CI)	Number of children (number of studies)
All	All	FE	2.1 (2.1, 2.2)	-20.6 (-20.7, −20.6)	24.9 (24.8, 24.9)	625618 (35)	61.5 (61.4, 61.6)	619100 (58)
		RE	-2.1 (-2.1, −2.0)	−25.0 (-25.1, −25.0)	20.9 (20.8, 21.0)		53.7 (53.6, 53.8)	
	<10 kg	FE	−4.4 (-4.5, −4.2)	−29.3 (-29.6, −29.0)	20.5 (20.3, 20.8)	26327 (13)	55.4 (54.8, 56.0)	27345 (19)
		RE	−1.7 (-1.9, −1.5)	−28.8 (-29.1, −28.5)	25.4 (25.1, 25.7)		54.9 (54.4, 55.5)	
	10–25 kg	FE	−3.8 (-3.9, −3.7)	−22.8 (-22.9, −22.6)	15.1 (14.9, 15.2)	51915 (15)	66.5 (66.1, 66.9)	58182 (21)
		RE	−1.6 (-1.7, −1.5)	−24.0 (-24.1, −23.8)	20.7 (20.5, 20.9)		61.7 (61.3, 62.1)	
	>25 kg	FE	−5.3 (-5.4, −5.3)	−29.4 (-29.6, −29.2)	18.7 (18.6, 18.9)	72803 (14)	53.6 (53.3, 54.0)	92969 (20)
		RE	−3.8 (-3.9, −3.7)	−34.7 (-35.0, −27.3)	27.1 (26.9, 27.3)		49.3 (49.0, 49.6)	
Category I (High income, low prevalence of underweight)	All	FE	−4.8 (-4.9, −4.7)	−27.8 (-27.9, −27.7)	18.2 (18.1, 18.3)	158926 (24)	60.1 (59.9, 60.2)	296662 (40)
		RE	−4.6 (-4.7, −4.5)	−30.2 (-30.3, −30.0)	20.9 (20.8, 21.1)		54.3 (54.1, 54.5)	
	<10 kg	FE	−4.5 (-4.7, −4.2)	−29.2 (-29.5, −28.9)	20.2 (19.9, 20.4)	25723 (9)	55.4 (54.8, 56.0)	26630 (14)
		RE	−3.3 (-3.5, −3.2)	−30.2 (-30.5, −29.9)	23.5 (23.2, 23.8)		54.2 (53.6, 54.8)	
	10–25 kg	FE	−4.0 (-4.1, −3.9)	−22.4 (-22.5, −22.2)	14.4 (14.2, 14.6)	49344 (10)	66.6 (66.2, 67.0)	55239 (15)
		RE	−3.4 (-3.5, −3.3)	−25.5 (-25.7, −25.3)	18.6 (18.5, 18.8)		62.9 (62.5, 63.3)	
	>25 kg	FE	−5.5 (-5.6, −5.4)	−29.3 (-29.5, −29.2)	18.3 (18.1, 18.4)	71030 (9)	53.5 (53.2, 53.9)	91092 (14)
		RE	−7.3 (-7.4, −7.1)	−39.6 (-39.9, −39.4)	25.1 (24.9, 25.3)		48.8 (48.5, 49.1)	
Category II (Middle income, medium prevalence of underweight)	All	FE	4.5 (4.4, 4.5)	−16.1 (-16.2, −16.1)	25.1 (25.0, 25.1)	462627 (7)	62.7 (62.6, 62.9)	465702 (10)
		RE	0.2 (0.1, 0.2)	−29.2 (-29.3, −29.1)	29.5 (29.4, 29.6)		56.3 (56.2, 56.4)	
	<10 kg	FE	−0.5 (-2.4, 1.4)	−35.9 (-39.4, −32.3)	34.9 (31.4, 38.4)	356 (2)	54.6 (50.1, 59.1)	467 (3)
		RE	−1.0 (-2.6, 0.7)	−31.6 (-34.7, −28.6)	29.7 (26.7, 32.8)		57.5 (53.0, 62.0)	
	10–25 kg	FE	−2.5 (-3.0, −1.9)	−28.5 (-29.5, −27.5)	23.5 (22.5, 24.6)	2246 (3)	63.2 (61.4, 65.1)	2618 (4)
		RE	−2.7 (-3.2, −2.2)	−26.3 (-27.2, −25.3)	20.9 (20.0, 21.9)		63.3 (61.4, 65.1)	
	>25 kg	FE	0.6 (−0.1, 1.4)	−29.5 (-29.7, −26.9)	29.5 (28.1, 30.9)	1498 (3)	61.3 (58.9, 63.7)	1602 (4)
		RE	−1.5 (-2.2, −0.8)	−29.2 (-30.5, −27.8)	26.2 (24.9, 27.6)		54.7 (52.3, 57.1)	
Category III (Low income, high prevalence of underweight)	All	FE	6.6 (6.2, 7.1)	−20.4 (-21.3, −19.6)	33.7 (32.8, 34.6)	3270 (6)	57.7 (57.1, 58.4)	19611 (9)
		RE	7.4 (6.9, 7.8)	−18.9 (-19.8, −18.1)	33.7 (32.8, 34.5)		42.2 (41.5, 42.9)	
	<10 kg	FE	3.8 (2.3, 5.4)	−21.2 (-24.2, −18.2)	28.9 (25.9, 31.8)	248 (2)	54.8 (48.6, 61.0)	248 (2)
		RE	4.9 (3.3, 6.4)	−19.7 (-22.7, −16.8)	29.4 (26.5, 32.4)		56.3 (50.1, 62.5)	
	10–25 kg	FE	9.3 (8.1, 10.6)	−12.8 (-15.1, −10.5)	31.5 (29.2, 33.8)	325 (2)	48.0 (42.6, 53.4)	325 (2)
		RE	8.8 (7.6, 10.0)	−12.9 (-15.1, −10.6)	30.5 (28.2, 32.7)		48.9 (43.5, 54.3)	
	>25 kg	FE	9.4 (7.6, 11.3)	−20.5 (-23.9, −17.1)	39.4 (36.0, 42.8)	275 (2)	39.6 (33.8, 45.4)	275 (2)
		RE	8.1 (6.3, 9.9)	−21.1 (-24.4, −17.8)	37.3 (34.0, 40.6)		42.0 (36.2, 47.8)	

dren of the same length. If habitus is not taken into account in the weight estimation, as has been recommended by previous authors, this is inevitable [2,5,16,17,32,44,55,58,73,81]. The pooled data showed an overall bias to underestimate weight, but with estimates as much as 30% below actual weight to 20% above actual weight. The differences between the populations with a different prevalence of underweight children were predictable: the Broselow tape underestimated the weight of fatter populations, and overestimated the weight of thinner populations. The weight estimation performance of the tape was not satisfactory in any population group.

The accuracy data showed a wide variation between studies, but virtually all studies failed to achieve satisfactory accuracy. Interestingly, the pooled results from populations with a medium prevalence of underweight children had the best accuracy results, most likely because of the lower prevalence of both obese and underweight children in these populations. Previous studies have shown the Broselow tape to be very accurate when limited to children of "normal" habitus, but the within- and between-population variability of body habitus precludes a consistently accurate performance, especially if the extremes of underweight or obesity are prevalent [58,81]. The poorest accuracy of the Broselow tape was in populations with many underweight children, potentially the sickest and most vulnerable to medication errors and as a result there is a very real possibility that the Broselow tape may produce a dangerously high weight overestimation in children with the highest risk of harm. This is especially true of the 2011 edition A tape: this tape is more inaccurate than older versions of the tape in low- and middle income countries because its shifted weight-zones based on high-income country growth data [79]. While additional drug doses may be administered or titrated to effect in overweight children, it is difficult to "take back" part of a dose given to an underweight child [115].

The meta-analysis pair-wise comparisons between the Broselow tape and other weight estimation systems provided some critical insights. The Broselow tape was significantly superior to age-based formulas and it is indisputably clear that age-based weight estimation should no longer be used [44,63,65,116–119]. Parental estimates of weight and Broselow estimates were not significantly different, but, given a choice between the two, parental estimates should probably be used [82]. Most notably, the Broselow tape was significantly inferior to the Mercy method, the Wozniak method and the PAWPER tape (the modern dual length- and habitus-based systems). This is further convincing evidence that safe, consistent and reproducible weight estimation cannot be performed using length alone, but length plus a measure of habitus can be extremely accurate [119,120]. The meta-analysis confirmed that there are weight estimation systems that are substantially more accurate than the Broselow tape, which should probably be considered in preference to the Broselow tape.

The advanced life support guidelines produced by both the American Heart Association and the Advance Life Support Group in the United Kingdom recommend the use of length-based weight estimation and advocate the use of length-based tapes [11,121]. These guidelines potentially need to be reconsidered in favour of dual length- and habitus-based weight estimation.

4.3 Children "too tall for the tape"

One of the other problems with the Broselow tape identified by researchers is those children "too tall for the tape" [53]. Many studies have reported a substantial proportion of children (ranging from 7% to 40%) for whom the Broselow tape could not generate an estimated weight [5,18,41,44,47,50,53,83]. According to the Broselow tape instructions, these children should be treated as adults with respect to drug doses, but this is potentially dangerous, especially as these children could be a young as 9 or 10 years of age [53]. There is no validated mechanism to procure an estimated weight using the Broselow tape for these weight estimation "orphans".

4.4 Potential habitus-based modification of Broselow tape weight estimation

While a system that has a large bias but high precision can be recalibrated, a system with low precision cannot be modified without introducing an additional variable to increase the precision. Several studies have explored the possibility of recalibrating the Broselow tape in different populations, with no success [29,62,122,123]. While this changed the bias of the Broselow tape, the overall accuracy was not improved. Only two recent studies which used a habitus-based modification of Broselow tape weight showed a meaningful improvement in accuracy, corroborating that habitus is the most important variable that affects weight-for-length [81,84]. The manufacturer's advocated method of modifying the weight estimation produced by the Broselow tape in obese children by shifting up a colour zone, has not yet been formally evaluated [102]. This system is also limited as it would not protect underweight children from overestimation of weight.

4.5 The Broselow tape and ideal body weight

The developers of the Broselow tape have defended its susceptibility to underestimate weight in populations with a high prevalence of obesity by suggesting that it actually predicts ideal body weight and not total body weight. They suggest that drug doses will therefore still be accurate in overweight and obese patients and that potential overdosing can be avoided, while underweight children should be tolerant to overdosing [102,124]. These claims are controversial and not yet established. Total body weight should still be targeted for drug dose calculations for children within the 95% BMI-for-age centile, until further definitive evidence emerges [11,105]. The optimum dosing for obese children (≥95% BMI-for-age centile) has yet to be established.

4.6 The Broselow tape and drug dose guidance

There were three main findings with respect to the Broselow tape's functioning as a drug dosing instrument. Firstly, when the tape was used as a stand-alone drug dosing guide, it was faster than standard drug reference systems (e.g. a textbook with drug doses) and produced fewer medication errors. But the occurrence of dosing errors was extremely high with the tape, producing dosing errors of more than 20% in up to 50% of dose calculations [10,89,95,100,107,113]. When drug dose determinations using the tape were compared with alternative drug-dosing systems that required no calculations to determine the appropriate drug volume to administer, the Broselow tape was inferior in both speed and accuracy to these systems [10,91,103,113]. Similarly, another tape method, the Handtevy system, because of its designated companion drug-dosing guide, produced fewer medication errors than the Broselow tape when used alone, although it estimated weight even less accurately than the Broselow tape [76,114]. The second major finding was that if the Broselow tape was used in conjunction with a supplementary colour-coded drug-dosing guide – a system that completely eliminated all calculations – then the frequency of dosing errors was reduced to less than 10% [93,100,106]. This was, perhaps, unsurprising given the success of other calculation-free methods. The third finding was that inadequate training with the tape was associated with poor performance, both in weight estimation as well as drug dose determination. Formal training reduced, but did not eliminate, errors, however [69,88,109].

Since the drug-dosing information on the Broselow tape is incomplete and the evidence is compelling that additional refer-

ence material must be used, the drug dose information on the tape must be concluded to be uncertain value. The statement by a major training organisation that, "it is reasonable to use a body length tape with pre-calculated doses" may perhaps be adapted to stipulate that the use of supplementary reference material that eliminates calculations is essential [11].

5 Limitations

The limitations of this study are similar to what may be encountered in any meta-analysis [13,125]. The heterogeneity between samples generally makes meta-analysis difficult, but this was countered by the use of subgroup analysis. Furthermore, the between-study variance was similar to the within-study variance which suggested that the pooled data analysis was valid. The significant difference in outcomes between the weight estimation systems was the final validation of the power of the systems to deal with this variance.

The incomplete reporting of data and the generally poor quality of data presentation and statistical analysis in many articles was the major limitation for this study. The relative paucity of accuracy data and direct, paired comparisons between the Broselow tape and other systems may have restricted the breadth of findings to some degree.

6 Conclusions

Medication errors in children are common, occur more often in the Emergency Department, in sicker children and when treated by doctors with little experience in managing emergencies in children [126,127]. Nearly one hundred articles have been published on the Broselow tape over the last three decades, which have provided data for the following conclusions.

The Broselow tape does not estimate weight as accurately as the newer length- and habitus-based systems such as the Mercy method and the PAWPER tape. Length-based weight estimation, although significantly superior to age-based weight estimation, is simply not accurate enough. The Broselow tape is also significantly limited by its length, with no proven method for dealing with children "too tall for the tape".

The Broselow tape on its own does not provide sufficient drug-dosing information to be used without additional, comprehensive colour-coded drug administration information. The calculations needed to translate the milligram doses on the tape to volumes of drug dilution to administer create a major, debilitating source of errors.

The Broselow tape has also been shown to be vulnerable to user-errors in under-trained individuals. It would be beneficial to explore the role of incorporating training in weight estimation into simulation exercises as part of advanced life support courses

Conflict of interest

Professor Wells developed the PAWPER tape, but from which he derives no financial benefit.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.resuscitation.2017.09.026.

References

[1]. Deboer S, Seaver M, Broselow J. Color coding to reduce errors. Am J Nurs 2005;105(8):68–71.

[2]. Lubitz DS, Seidel JS, Chameides L, Luten RC, Zaritsky AL, Campbell FW. A rapid method for estimating weight and resuscitation drug dosages from length in the pediatric age group. Ann Emerg Med 1988;17(6):576–81, http://dx.doi.org/10.1016/s0196-0644(88)80396-2.

[3]. Hughes G, Spoudeas H, Kovar IZ, Millington HT. Tape measure to aid prescription in paediatric resuscitation. Arch Emerg Med 1990;7(1):21–7.

[4]. Milne K, Yasin A, Knight J, Noel D, Lubell R, Filler G. Ontario children have outgrown the Broselow tape. Can J Emerg Med 2012;14(1):25–30.

[5]. Nieman CT, Manacci CF, Super DM, Mancuso C, Fallon Jr WF. Use of the Broselow tape may result in the underresuscitation of children. Acad Emerg Med 2006;13(10):1011–9, http://dx.doi.org/10.1197/j.aem.2006.06.042.

[6]. Clark MC, Lewis RJ, Fleischman RJ, Ogunniyi AA, Patel DS, Donaldson RI. Accuracy of the broselow tape in south Sudan, the hungriest place on earth. Acad Emerg Med 2016;23(1):21–8, http://dx.doi.org/10.1111/acem.12854.

[7]. Dicko A, Alhousseini ML, Sidibé B, Traoré M, Abdel-Rahman SM. Evaluation of the Mercy weight estimation method in Ouelessebougou, Mali. BMC Public Health 2014;14:270, http://dx.doi.org/10.1186/1471-2458-14-270.

[8]. Batmanabane G, Kumar Jena P, Dikshit R, Abdel-Rahman SM. Using the Mercy Method for weight estimation in Indian children. Global Pediatric Health 2015;(January–December):1–7.

[9]. Hohenhaus SM, Frush KS. Pediatric patient safety: common problems in the use of resuscitative aids for simplifying pediatric emergency care. J Emerg Nurs 2004;30(1):49–51, http://dx.doi.org/10.1016/j.jen.2003.11.011.

[10]. Moreira ME, Hernandez C, Stevens AD, Jones S, Sande M, Blumen JR, et al. Color-coded prefilled medication syringes decrease time to delivery and dosing error in simulated Emergency Department pediatric resuscitations. Ann Emerg Med 2015, http://dx.doi.org/10.1016/j.annemergmed.2014.12.035.

[11]. Kleinman ME, Chameides L, Schexnayder SM, Samson RA, Hazinski MF, Atkins DL, et al. Pediatric advanced life support: 2010 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Pediatrics 2010;126(5):e1361–99, http://dx.doi.org/10.1542/peds.2010-2972D.

[12]. Waseem M, Chen J, Leber M, Giambrone AE, Gerber LM. A reexamination of the accuracy of the Broselow Tape as an instrument for weight estimation. Pediatr Emerg Care 2017, http://dx.doi.org/10.1097/pec.0000000000000982.

[13]. Williamson PR, Lancaster GA, Craig JV, Smyth RL. Meta-analysis of method comparison studies. Stat Med 2002;21(14):2013–25, http://dx.doi.org/10.1002/sim.1158.

[14]. Global Burden of Disease Study 2013. Global Burden of Disease Study 2013 (GBD 2013) Obesity Prevalence 1990–2013. Seattle, United States: Institute for Health Metrics and Evaluation (IHME); 2014.

[15]. Dearlove CB, Dearlove O. Visual estimation of children's weights. Anaesthesia 1999;54(12):1228–9, http://dx.doi.org/10.1046/j.1365-2044.1999.01229.x.

[16]. Kun W, Cheng K, Yuen M, Tung W. How good is the Broselow tape measurement for estimation of body weights in paediatric patients for application in Hong Kong? Hong Kong J Paediatr 2000;5(1):25–30.

[17]. Black K, Barnett P, Wolfe R, Young S. Are methods used to estimate weight in children accurate? Emerg Med 2002;14(2):160–5, http://dx.doi.org/10.1046/j.1442-2026.2002.00311.x.

[18]. Hofer CK, Ganter M, Tucci M, Klaghofer R, Zollinger A. How reliable is length-based determination of body weight and tracheal tube size in the paediatric age group? The Broselow tape reconsidered. Br J Anaesth 2002;88(2):283–5.

[19]. Argall JA, Wright N, Mackway-Jones K, Jackson R. A comparison of two commonly used methods of weight estimation. Arch Dis Childhood 2003;88(9):789–90.

[20]. Potier K. Broselow tape or APLS formula to estimate weight in children. Emerg Med J 2006.

[21]. Varghese A, Vasudevan VK, Lewin S, Indumathi CK, Dinakar C, Subba Rao SD. Do the length-based (Broselow®) tape, APLS, Argall and Nelson's formulae accurately estimate weight of Indian children? Indian Pediatr 2006;43(10):889–94.

[22]. DuBois D, Baldwin S, King WD. Accuracy of weight estimation methods for children. Pediatr Emerg Care 2007;23(4):227–30, http://dx.doi.org/10.1097/PEC.0b013e31803f5aca.

[23]. Hashikawa A, Juhn Y, Homme J, Gardner B, Moore B. Does length-based resuscitation tape accurately place pediatric patients into appropriate color-coded zones? Pediatr Emerg Care 2007;23(12):856–61.

[24]. Im T, Kang B, Choi H, Won T, Yeon J. A study on the accuracy of weight estimation using the Broselow tape in Korean children. Ann Emerg Med 2007;50(3):S40.

[25]. Jang HY, Shin SD, Kwak YH. Can the Broselow tape be used to estimate weight and endotracheal tube size in Korean children? Acad Emerg Med 2007;14(5):489–91.

[26]. Krieser D, Nguyen K, Kerr D, Jolley D, Clooney M, Kelly AM. Parental weight estimation of their child's weight is more accurate than other weight estimation methods for determining children's weight in an emergency department? Emerg Med J 2007;24(11):756–9, http://dx.doi.org/10.1136/emj.2007.047993.

[27]. Patel A, Zikos V, Rzechula K, McQuillen K. The accuracy of pediatric weight in the Emergency Department: do caregiver weight estimates measure up? Ann Emerg Med 2007;50(3 Supplement):S123–4, http://dx.doi.org/10.1016/j.annemergmed.2007.06.443.

[28]. Zikos V, Patel A, Rzechula KKM. The accuracy of pediatric weight in the emergency department: do caregiver weight estimates measure up? Ann Emerg Med 2007;50(3):S123–4.

[29]. Ramarajan N, Krishnamoorthi R, Strehlow M, Quinn J, Mahadevan SV. Internationalizing the Broselow tape: how reliable is weight estimation in Indian children. Acad Emerg Med 2008;15(5):431–6, http://dx.doi.org/10.1111/j.1553-2712.2008.00081.x.

[30]. Zink K, Bohn M, Hood T, Topp RV, Berger J. A comparison of methods of assessing patient body weight in the pediatric emergency department. Ky Nurse 2008;56(3):12.

[31]. Anstett D, Bawden J, Moylette E. Does the Broselow tape accurately estimate the weight of healthy Irish children? Can J Emerg Med 2009;11:289.

[32]. Cattamanchi S, Banala S, Trichur SV. Assessing competency of the Broselow-Luten pediatric resuscitation tape: a prospective, cross sectional, analytical study of 15, 000 Indian school children. Ann Emerg Med 2009;54(3):S130–1.

[33]. Cattermole GN, Leung M, Mak P, So HK, Graham CA, Rainer TH. Children's weights correlate more strongly with mid-arm circumference (MAC) than with age, height or foot-length. J Emerg Med 2009;37(2):228–9, http://dx.doi.org/10.1016/j.jemermed.2009.06.051.

[34]. So TY, Farrington E, Absher RK. Evaluation of the accuracy of different methods used to estimate weights in the pediatric population. Pediatrics 2009;123(6):e1045–51, http://dx.doi.org/10.1542/peds.2008-1968.

[35]. Stewart D. Accuracy of the Broselow tape for estimating paediatric weight in two Australian Emergency Departments. University of Sydney; 2009. Available from: https://ses.library.usyd.edu.au//bitstream/2123/6265/1/Declan Stewart WPP final.pdf.

[36]. Yamamoto LG, Inaba AS, Young LL, Anderson KM. Improving length-based weight estimates by adding a body habitus (obesity) icon. Am J Emerg Med 2009;27(7):810–5, http://dx.doi.org/10.1016/j.ajem.2008.06.023.

[37]. Casey J, Borland M. Best Guess method: a further external validation study and comparison with other methods. Emerg Med Aust 2010;22(1):68–74, http://dx.doi.org/10.1111/j.1742-6723.2009.01258.x.

[38]. Cattermole GN, Leung PY, Mak PS, Graham CA, Rainer TH. Mid-arm circumference can be used to estimate children's weights. Resuscitation 2010;81(9):1105–10, http://dx.doi.org/10.1016/j.resuscitation.2010.05.015.

[39]. Lulic I, Kovic I. Comparing methods for weight estimation of children: a preliminary study. Resuscitation 2010;81(2 Supplement):S77, http://dx.doi.org/10.1016/j.resuscitation.2010.09.315.

[40]. Bourdeau S, Copeland J, Milne WK. Accuracy of the Broselow tape in estimating the weight of First Nations children. Can J Rural Med 2011;16(4):121–5.

[41]. Geduld H, Hodkinson PW, Wallis LA. Validation of weight estimation by age and length based methods in the Western Cape, South Africa population. Emerg Med J 2011;28(10):856–60, http://dx.doi.org/10.1136/emj.2010.098640.

[42]. Knight JC, Nazim M, Riggs D, Channel J, Mullet C, Vaughan R, et al. Is the Broselow tape a reliable indicator for use in all pediatric trauma patients?: a look at a rural trauma center. Pediatr Emerg Care 2011;27(6):479–82, http://dx.doi.org/10.1097/pec.0b013e31821d8559.

[43]. Rosenberg M, Greenberger S, Rawal A, Latimer-Pierson J, Thundiyil J. Comparison of Broselow tape measurements versus physician estimations of pediatric weights. Am J Emerg Med 2011;29(5):482–8, http://dx.doi.org/10.1016/j.ajem.2009.12.002.

[44]. Abdel-Rahman SM, Ridge AL. An improved pediatric weight estimation strategy. Open Med Dev J 2012;4:87–97, http://dx.doi.org/10.2174/1875181401204010087.

[45]. Heyming T, Bosson N, Kurobe A, Kaji AH, Gausche-Hill M. Accuracy of paramedic Broselow tape use in the prehospital setting. Prehospital Emerg Care 2012;16(3):374–80, http://dx.doi.org/10.3109/10903127.2012.664247.

[46]. Park J, Kwak YH, Kim do K, Jung JY, Lee JH, Jang HY, et al. A new age-based formula for estimating weight of Korean children. Resuscitation 2012;83(9):1129–34, http://dx.doi.org/10.1016/j.resuscitation.2012.01.023.

[47]. Sinha M, Lezine MW, Frechette A, Foster KN. Weighing the pediatric patient during trauma resuscitation and its concordance with estimated weight using Broselow Luten emergency tape. Pediatr Emerg Care 2012;28(6):544–7, http://dx.doi.org/10.1097/pec.0b013e318258ac2e.

[48]. Trakulsrichai S, Boonsri C, Chatchaipun P, Chunharas A. Accuracy of three methods used for Thai children's body weight estimation. J Med Assoc Thailand 2012;95(9):1194–9.

[49]. Wozniak R. The evaluation of potential weight-estimation methods in a primarily HIV positive cohort in Botswana for use in resource limited settings. The University of British Columbia; 2012. Available from:.

[50]. Abdel-Rahman SM, Paul IM, James LP, Lewandowski A. Evaluation of the Mercy TAPE: performance against the standard for pediatric weight estimation. Ann Emerg Med 2013;62(4):332–9, http://dx.doi.org/10.1016/j.annemergmed.2013.02.021, e6.

[51]. Abdel-Rahman SM, Ahlers N, Holmes A, Wright K, Harris A, Weigel J, et al. Validation of an improved pediatric weight estimation strategy. Journal of Pediatr Pharmacology and Therapeutics 2013;18(2):112–21, http://dx.doi.org/10.5863/1551-6776-18.2.112.

[52]. Akabarian S, Vahedparast H, Ravanipour M, Mirzaei K. Comparison of Broselow tape measurements versus mother estimations of pediatric weights. Iran South Med J 2013;16(2):138–43.

[53]. Cattermole GN, Leung PY, Graham CA, Rainer TH. Too tall for the tape: the weight of schoolchildren who do not fit the Broselow tape. Emerg Med J 2013, http://dx.doi.org/10.1136/emermed-2012-202325 [epub ahead of print].

[54]. Graves L, Chayen G, Peat J, O'Leary F. A comparison of actual to estimated weights in Australian children attending a tertiary children's hospital, using the original and updated APLS, Luscombe and Owens, Best Guess formulae and the Broselow tape. Resuscitation 2014;85(3):392–6, http://dx.doi.org/10.1016/j.resuscitation.2013.11.024.

[55]. House DR, Ngetich E, Vreeman RC, Rusyniak DE. Estimating the weight of children in Kenya: do the Broselow tape and age-based formulas measure up? Ann Emerg Med 2013;61(1):1–8, http://dx.doi.org/10.1016/j.annemergmed.2012.07.110.

[56]. Loo PY, Chong SL, Lek N, Bautista D, Ng KC. Evaluation of three paediatric weight estimation methods in Singapore. J Paediatr Child Health 2013;49(4):E311–6, http://dx.doi.org/10.1111/jpc.12141.

[57]. Suh D, Kwak Y, Kim D, Jeong J. Estimating weights using the Broselow tape in Korea: comparison of old and new versions of the Broselow tape use. Ann Emerg Med 2013;62(4):S109.

[58]. Wells M, Coovadia A, Kramer E, Goldstein L. The PAWPER tape: a new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. Resuscitation 2013;84(2):227–32, http://dx.doi.org/10.1016/j.resuscitation.2012.05.028.

[59]. Young TP, Chen BG, Kim TY, Thorp AW, Brown L. Finger counting: an alternative method for estimating pediatric weights. Am J Emerg Med 2014;32(3):243–7, http://dx.doi.org/10.1016/j.ajem.2013.11.034.

[60]. Allison N, Norton I. Emergency weight estimation in Aboriginal and Torres Strait Islander children in the Northern Territory: are the current methods accurate? Emerg Med Australas 2014;26(5):487–93, http://dx.doi.org/10.1111/1742-6723.12278.

[61]. Chiengkriwate P, Donnapee R, Geater A. The accuracy of the Broselow tape in the weight estimation of Thai children. Asian Biomed 2014;8(6):799–807.

[62]. Asskaryar F, Shankar R. An Indian pediatric emergency weight estimation tool: prospective adjustment of the Broselow tape. Int J Emerg Med 2015;8(1):78, http://dx.doi.org/10.1186/s12245-015-0078-z.

[63]. Britnell S, Koziol-McLain J. Weight estimation in paediatric resuscitation: a hefty issue in New Zealand. Emerg Med Australas 2015;27(3):251–6, http://dx.doi.org/10.1111/1742-6723.12389.

[64]. Chavez H, Peterson R, Lo K, Arel M. Weight estimation in an inner city Pediatric Emergency Department: the effect of obesity. Am J Emerg Med 2015;33(10):1364–7, http://dx.doi.org/10.1016/j.ajem.2015.07.040.

[65]. Georgoulas V, Wells M. The PAWPER tape and the Mercy Method outperform other methods of weight estimation in children in South Africa. South Afr Med J 2016;106(9):933–9, http://dx.doi.org/10.7196/SAMJ.2016.v106i9.10572.

[66]. Khouli M, Ortiz MI, Romo-Hernández G, Martínez-Licona D, Stelzner SM. Use of the broselow tape in a Mexican Emergency Department. J Emerg Med 2015;48(6):660–6, http://dx.doi.org/10.1016/j.jemermed.2014.12.082.

[67]. O'Leary F, John-Denny B, McGarvey K, Hann A, Pegiazoglou I, Peat J. Estimating the weight of ethnically diverse children attending an Australian emergency department: a prospective, blinded, comparison of age-based and length-based tools including Mercy, PAWPER and Broselow. Arch Dis Childhood 2017;102(1):46–52, http://dx.doi.org/10.1136/archdischild-2016-310917.

[68]. Young TP, Washington O, Flanery A, Guptill M, Reibling ET, Brown L, et al. Comparison of the finger counting method, the Broselow tape and common weight estimation formulae in Filipino children after Typhoon Haiyan. Emerg Med Australas 2015;27(3):239–44, http://dx.doi.org/10.1111/1742-6723.12382.

[69]. Abdel-Rahman A, Jacobsen R, Watts J, Doyle S, O'Malley D, Hefner T, et al. Comparative performance of pediatric weight estimation techniques: a human factor errors analysis. Pediatr Emerg Care 2015, http://dx.doi.org/10.1097/pec.0000000000000543.

[70]. Talib NJ, Rahm G, Abdel-Rahman SM. Estimating weight in children with Down Syndrome. Global Pediatr Health 2015;2, http://dx.doi.org/10.1177/2333794x14568450.

[71]. AlHarbi TM, AlGarni A, AlGamdi F, Jawish M, Wani TA, Abu-Shaheen AK. The accuracy of Broselow Tape weight estimate among pediatric population. Biomed Res Int 2016;2016:7042947, http://dx.doi.org/10.1155/2016/7042947.

[72]. Aliyu I, Abdulsalam M. Is the broselow tape and APLS formula applicable to nigerian children? The Kano experience. Trop J Med Res 2016;19(1):20–3.

[73]. Bowen L, Zyambo M, Snell D, Kinnear J, Bould MD. Evaluation of the accuracy of common weight estimation formulae in a Zambian paediatric surgical population. Anaesthesia 2017;72(4):470–8, http://dx.doi.org/10.1111/anae.13780.

[74]. Jung JY, Kwak YH, Kim DK, Suh D, Chang I, Yoon C, et al. An assessment of the accuracy of a novel weight estimation device for children. Emerg Med J 2016, http://dx.doi.org/10.1136/emermed-2016-205817.

[75]. Jung J, Kwon H, Choi Y. Methods of weight estimation in pediatric resuscitation. Pediatr Emerg Med J 2016;3(1):9–14.

[76]. Lowe CG, Campwala RT, Ziv N, Wang VJ. The Broselow and Handtevy Resuscitation Tapes: a comparison of the performance of pediatric weight prediction. Prehospital Disaster Med 2016;31(4):364–75, http://dx.doi.org/10.1017/S1049023x16000455.

[77]. Mishra DG, Kole T, Nagpal R, Smith JP. A correlation analysis of Broselow Pediatric Emergency Tape-determined pediatric weight with actual pediatric weight in India. World J Emerg Med 2016;7(1):40–3, http://dx.doi.org/10.5847/wjem.j.1920-8642.2016.01.007.

[78]. Nosaka N, Fujiwara T, Knaup E, Okada A, Tsukahara H. Validity of mothers' reports of children's weight in Japan. Acta Medica Okayama 2016;70(4):255–9.

[79]. Ralston ME, Myatt MA. Weight estimation tool for children aged 6–59 months in limited-resource settings. PLoS One 2016;11(8):e0159260, http://dx.doi.org/10.1371/journal.pone.0159260.

[80]. Sahar M, Nordin N, Nor A, Zakaria M. Weight estimation method in emergency department in Malaysia: is broselow tape (BT) reliable? Jurnal Teknologi (Sciences & Engineering) 2016;78(6–8):113–7.
[81]. Tanner D, Negaard A, Huang R, Evans N, Hennes H. A prospective evaluation of the accuracy of weight estimation using the Broselow Tape in overweight and obese pediatric patients in the Emergency Department. Pediatr Emerg Care 2016, http://dx.doi.org/10.1097/pec.0000000000000894.
[82]. Young KD, Korotzer NC. Weight estimation methods in children: a systematic review. Ann Emerg Med 2016;68(4):441–51.e10, http://dx.doi.org/10.1016/j.annemergmed.2016.02.043, e10.
[83]. Wells M, Goldstein L, Bentley A. A validation study of the PAWPER XL tape: accurate estimation of both total and ideal body weight in children up to 16 years of age. Trauma Emerg Care 2017;2(4):1–8, http://dx.doi.org/10.15761/TEC.1000141.
[84]. Wells M, Goldstein L, Bentley A. Accuracy of weight estimation by the Broselow tape is substantially improved by including a visual assessment of body habitus. Pediatric Res 2017, http://dx.doi.org/10.1038/pr.2017.222 (in press).
[85]. Whitfield KC, Wozniak R, Pradinuk M, Karakochuk CD, Anabwani G, Daly Z, et al. Anthropometric measures are simple and accurate paediatric weight-prediction proxies in resource-poor settings with a high HIV prevalence. Arch Dis Child 2017;102(1):10–6, http://dx.doi.org/10.1136/archdischild-2015-309645.
[86]. Luten RC, Wears RL, Broselow J, Zaritsky A, Barnett TM, Lee T, et al. Length-based endotracheal tube and emergency equipment in pediatrics. Ann Emerg Med 1992;21(8):900–4, http://dx.doi.org/10.1016/s0196-0644(05)82924-5.
[87]. Lanoix R, Golden J. The facilitated pediatric resuscitation room. J Emerg Med 1999;17(2):363–6.
[88]. Hohenhaus SM. Is this a drill? Improving pediatric emergency preparedness in North Carolina's emergency departments. J Emerg Nurs 2001;27(6):568–70, http://dx.doi.org/10.1067/men.2001.119686.
[89]. Vilke GM, Marino A, Fisher R, Chan TC. Estimation of pediatric patient weight by EMT-PS. J Emerg Med 2001;21(2):125–8, http://dx.doi.org/10.1016/s0736-4679(01)00354-7.
[90]. Luten R, Wears RL, Broselow J, Croskerry P, Joseph MM, Frush K. Managing the unique size-related issues of pediatric resuscitation: reducing cognitive load with resuscitation aids. Acad Emerg Med 2002;9(8):840–7, http://dx.doi.org/10.1197/aemj.9.8.840.
[91]. Slishman S, Sapien R, Crandall CS. Introducing a simple, weight-based, color-coded, medication dosing device. Pediatr Emerg Care 2002;18(3):212–5.
[92]. McDaniel-Hohenhaus S, Frush K. Development of an e-learning program for use of the Broselow-Luten pediatric system. J Emerg Nurs 2003;29(5):417.
[93]. Shah AN, Frush K, Luo X, Wears RL. Effect of an intervention standardization system on pediatric dosing and equipment size determination: a crossover trial involving simulated resuscitation events. Arch Pediatr Adolesc Med 2003;157(3):229–36.
[94]. DeBoer SL, Seaver M, Boatright JR, Robinson K. Pediatric spinal immobilization: c-spines, car seats, and color-coded collars. J Emerg Nurs 2004;30(5):481–4.
[95]. Moore B, Hankins D, Shimmin S. Use of the length-based resuscitation tape for estimating pediatric weight in the pre-hospital setting. Prehospital Emerg Care 2004;8(1):87.
[96]. Agarwal S, Swanson S, Murphy A, Yaeger K, Sharek P, Halamek LP. Comparing the utility of a standard pediatric resuscitation cart with a pediatric resuscitation cart based on the Broselow tape: a randomized, controlled, crossover trial involving simulated resuscitation scenarios. Pediatrics 2005;116(3):e326–33, http://dx.doi.org/10.1542/peds.2005-0320.
[97]. DeBoer S, Seaver M, Broselow J. Do you know your ABCs? Airway, breathing, and colour-coding! Australas Emerg Nurs J 2005;8(1-2):35–41, http://dx.doi.org/10.1016/j.aenj.2005.04.006.
[98]. Hohenhaus SM, Frush KS. Pediatric patient safety in resuscitation simulations: lessons learned from studying clinician behavior. Pediatr Emerg Care 2005;21(1):71–5.
[99]. Frush K, Hohenhaus S, Luo X, Gerardi M, Wiebe RA. Evaluation of a Web-based education program on reducing medication dosing error: a multicenter, randomized controlled trial. Pediatr Emerg Care 2006;22(1):62–70.
[100]. Kaji AH, Gausche-Hill M, Conrad H, Young KD, Koenig WJ, Dorsey E, et al. Emergency medical services system changes reduce pediatric epinephrine dosing errors in the prehospital setting. Pediatrics 2006;118(4):1493–500, http://dx.doi.org/10.1542/peds.2006-0854.
[101]. Frush KS. Medication errors in pediatric emergency care: developing a national standard? Ann Emerg Med 2007;50(4):369–70.
[102]. Luten RC, Zaritsky A, Wears R, Broselow J. The use of the Broselow tape in pediatric resuscitation. Acad Emerg Med 2007;14(5):500–1, http://dx.doi.org/10.1197/j.aem.2007.02.015, author reply 501-2.
[103]. Fineberg SL, Arendts G. Comparison of two methods of pediatric resuscitation and critical care management. Ann Emerg Med 2008;52(1):35–40, http://dx.doi.org/10.1016/j.annemergmed.2007.10.021, e13.
[104]. Luten RC, Broselow J. Standardization of product concentration in emergency dosing: a response to Fineberg and Arendt. Ann Emerg Med 2008;52(4):477–8, http://dx.doi.org/10.1016/j.annemergmed.2008.05.038, author reply 478.
[105]. Wells M, Kramer E. Optimizing emergency drug dosing in children. Acad Emerg Med 2008;15(12):1325.
[106]. Feleke R, Kalynych CJ, Lundblom B, Wears R, Luten R, Kling D. Color coded medication safety system reduces community pediatric emergency nursing medication errors. J Patient Saf 2009;5(2):79–85, http://dx.doi.org/10.1097/pts.0b013e3181a647ab.
[107]. Pinchevsky LE, Pesaturo KA, Smith BS, Hartman CA. Pilot comparison of three cardiopulmonary resuscitation medication dosing strategies in overweight children. J Pediatr Pharmacol Ther 2010;15(4):282–9.
[108]. Kelly A. Estimation of children's weight in medical emergencies. In: Preedy V, editor. Handbook of anthropometry: Physical measures of human form in health and disease. New York: Springer-Verlag; 2012. p. 1151–61.
[109]. Lammers R, Byrwa M, Fales W. Root causes of errors in a simulated prehospital pediatric emergency. Acad Emerg Med 2012;19(1):37–47, http://dx.doi.org/10.1111/j.1553-2712.2011.01252.x.
[110]. Meguerdichian MJ, Clapper TC. The Broselow tape as an effective medication dosing instrument: a review of the literature. J Pediatr Nurs 2012;27(4):416–20, http://dx.doi.org/10.1016/j.pedn.2012.04.009.
[111]. Abdel-Rahman SM, Ridge A, Kearns GL. Estimation of body weight in children in the absence of scales: a necessary measurement to insure accurate drug dosing. Arch Dis Childhood 2014;99(6):570–4, http://dx.doi.org/10.1136/archdischild-2013-305211.
[112]. Lammers R, Willoughby-Byrwa M, Fales W. Medication errors in prehospital management of simulated pediatric anaphylaxis. Prehospital Emerg Care 2014;18(2):295–304, http://dx.doi.org/10.3109/10903127.2013.856501.
[113]. Stevens AD, Hernandez C, Jones S, Moreira ME, Blumen JR, Hopkins E, et al. Color-coded prefilled medication syringes decrease time to delivery and dosing errors in simulated prehospital pediatric resuscitations: a randomized crossover trial. Resuscitation 2015;96:85–91, http://dx.doi.org/10.1016/j.resuscitation.2015.07.035.
[114]. Rappaport LD, Brou L, Givens T, Mandt M, Balakas A, Roswell K, et al. Comparison of errors using two length-based tape systems for prehospital care in children. Prehosp Emerg Care 2016;20(4):508–17, http://dx.doi.org/10.3109/10903127.2015.1128027.
[115]. Wells M, Goldstein L, Botha M. Emergency drug dosing in children. Edinburgh: Churchill Livingstone Elsevier; 2012.
[116]. Marlow R, Lo D, Walton L. Accurate paediatric weight estimation by age: mission impossible? Arch Dis Child 2011;96(Suppl. 1):A1–2, http://dx.doi.org/10.1136/adc.2011.212563.3.
[117]. Cattermole GN, Leung MP, So HK, Mak PS, Graham CA, Rainer TH. Age-based formulae to estimate children's weight in the emergency department. Emerg Med J 2011;28(5):390–6, http://dx.doi.org/10.1136/emj.2009.090357.
[118]. Cattermole GN, Graham CA, Rainer TH. Age-based weight prediction is imprecise. Emerg Med J 2011;29(1):82–3.
[119]. Wells M, Goldstein L, Bentley A. It is time to abandon age-based emergency weight estimation in children! A failed validation of 20 different age-based formulas. Resuscitation 2017;116(7):73–83.
[120]. Losek JD, Garland J, Nelson DB. Body habitus + height = accurate weight estimate. Ann Emerg Med 1989;18(6):709, http://dx.doi.org/10.1016/s0196-0644(89)80539-6.
[121]. Advanced Life Support Group. Advanced Paediatric Life Support: The Practical Approach (APLS). 5th edition BMJ Books – Publisher: John Wiley & Sons (Wiley-Blackwell); 2011.
[122]. Sandell JM, Charman SC. Can age-based estimates of weight be safely used when resuscitating children? Emerg Med J 2009;26(1):43–7, http://dx.doi.org/10.1136/emj.2008.061119.
[123]. Theron L, Adams A, Jansen K, Robinson E. Emergency weight estimation in Pacific Island and Maori children who are large-for-age. Emerg Med Australas 2005;17:238–43.
[124]. Luten R, Zaritsky A. The sophistication of simplicity... optimizing emergency dosing. Acad Emerg Med 2008;15(5):461–5, http://dx.doi.org/10.1111/j.1553-2712.2008.00107.x.
[125]. Berman NG, Parker RA. Meta-analysis: neither quick nor easy. BMC Med Res Methodol 2002;2:10.
[126]. Kozer E, :maintitle>Meta-analysis: neither quick nor easyBMC Med Res Methodol2200210[126]E.Kozer D, Scolnik D, Macpherson A, Keays T, Shi K, Luk T, et al. Variables associated with medication errors in paediatric emergency medicine. Pediatrics 2002;110:737–42.
[127]. Kaufmann J, Laschat M, Wappler F. Medication errors in pediatric emergencies: a systematic analysis. Deutsches Ärzteblatt Int 2012;109(38):609–16, http://dx.doi.org/10.3238/arztebl.2012.0609.

Supplementary table 1 Modified Newcastle-Ottawa scale as used in this study.

Newcastle-Ottawa Scale adapted for method-comparison studies

Selection: (Maximum 5 stars)

1. Representativeness of the study sample:
    a. Truly representative of the target population. (all subjects or random sampling)□★
    b. Somewhat representative of the target population (non-random sampling)□★
    c. Selected group.
    d. No description of the sampling strategy.
2. Sample size:
    a. Justified and acceptable ★
    b. Not justified.
3. Prospective, non-database data ★
4. Appropriate application of weight-estimation technique ★

Comparability: (Max 2 stars) (all data is paired within each study.)

1. Paired data ★★

Outcome: (Maximum 3 stars)

1. Assessment of the outcome:
    a. Independent blinded assessment ★★
    b. Independent non-blinded assessment ★
    c. Non-blinded assessment ★
    d. No description

If data analysis was not appropriate for method-comparison studies (no assessment of bias, precision (or repeatability if applicable) and accuracy) then consider downgrading one level.

2. Statistical test:
    a. Appropriate, clearly described statistical analysis with confidence intervals or measure of variance ★
    b. The statistical test is not appropriate, not described or incomplete.

If study did not group ages / weight categories appropriately or alternatively use percentage error analysis or logarithmic transformation, then downgrade to ‘b’.

Supplementary table 2 Raw data from the included studies. The bold, red font indicates data that was imputed. The bold, blue font indicates accuracy data that was imputed from mean percentage error data.

Broselow Tape							
Country	**N**	**Subgroup**	**MPE [SD] (%)**	**PELOA (%)**	**PW10 (%)**	**PW20 (%)**	**Author**
USA 1988 BT 1998	937 395 449 93	All 3.5-10kg 10-25kg >25kg			59.7 55.9 65.0 49.5	89.6 85.6 95.1 83.9	Lubitz
UK 1999 *BT version NR*	50	All			60.0		Dearlove
Hong Kong 2000 *BT version NR*	909 129 631 149	All <10kg 10-25kg >25kg			69.5 56.6 76.2 52.3	91.3 83.7 95.2 81.2	Kun
Australia 2002 *BT version NR*	339 121 132 86	All <10kg 10-25kg 25-40kg	**-2.0 [11.7]** -0.6 **[12.3]** -0.4 **[9.4]** -6.4 **[12.8]**	**-24.9, 20.9** **-24.7, 23.5** **-18.8, 20.9** **-31.5, 18.7**	**60.0** **58.3** **71.2** **51.1**	**90.8** **89.6** **96.6** **83.6**	Black
Switzerland 2002 *BT version NR*	585 339 246	All ≤20kg >20kg			65.0 66.0 65.0		Hofer
New Zealand 2005 *BT version NR*	909 226 420 160 79	All Maori Pacific European Asian	**-9.5 [12.6]** -11.1 [13.6] -11.0 [11.0] -5.9 [12.5] -3.9 [14.0]	**-34.1, 15.1** **-37.9, 15.6** **-32.6, 10.5** **-30.4, 18.7** **-31.2, 23.5**	**45.5** **40.7** **43.6** **52.7** **50.8**	**78.8** **73.2** **79.1** **85.1** **83.1**	Theron
USA 2006 BT1998	7813 2305 2660 2848	All <10kg 10-20kg >20kg	-3.9 [11.9] **-3.6 [12.7]** **-5.0 [10.9]** **-3.8 [15.6]**	**-27.2, 19.4** **-28.4, 21.3** **-26.4, 16.5** **-34.4, 26.7**	55.3 57.0 59.8 53.2	86.6 **87.0** **90.5** **78.7**	Nieman
USA 2006 BT2005A	7671 2305 2660 2706	All <10kg 10-20kg >20kg	-5.6 [11.9] **-2.9 [11.9]** **-5.7 [9.5]** **-7.8 [13.4]**	**-28.9, 17.7** **-26.3, 20.5** **-24.4, 13.0** **-34.0, 18.5**	60.0 62.3 62.0 58.5	93.7 **89.7** **93.0** **80.0**	Nieman
USA 2007 *BT version NR*	300 100 100 100	All <10kg 10-20kg 20-36kg	**-8.2 [19.1]** -9.9 [15.5] -7.1 [11.3] -7.5 [26.9]	**-45.6, 29.3** **-40.3, 20.5** **-29.3, 15.0** **-60.1, 45.1**	**36.7** **40.3** **53.6** **27.9**	**66.2** **71.6** **86.5** **52.6**	DuBois
USA 2007 BT2002B	1207 544 520 143	All Birth 4-6yrs 10-12yrs			**42.1**	66.2 100 70.0 40.6	Hashikawa
Korea 2007 *BT version NR*	665	All ≤26kg >26kg			57.9 64.5 38.5		Jang
Australia 2007 *BT version NR*	410	All			61.0		Krieser
India 2008 *BT version NR*	548 175 197 176	All <10kg 10-18kg >18kg	**9.0 [13.1]** 2.4 [12.6] 11.3 [10.9] 12.9 [13.5]	**-16.8, 34.7** **-22.3, 27.1** **-10.1, 32.7** **-13.6, 39.4**	**43.6** 52.6 44.7 33.5	**78.6** **88.1** **78.6** **69.3**	Ramarajan
Ireland 2009 *BT version NR*	545	All	-8.6 [13.2]	-34.5, 17.2	69.7	86.6	Anstett
India 2009 *BT version NR*	15000 4256 6076 4668	All <10kg 10-18kg >18kg	**-11.6** -2.4 -13.3 -17.9		63.0 57.0		Cattamanchi
USA 2009 *BT version NR*	1011 471 382 119 39 166 650 195	All <10kg 10-25kg 25-40kg >40kg UW NW OW	**-4.2 [21.6]** 0.2 [24.0] -3.7 [16.0] -12.0 [17.0] -38.0 [12.0] 24.0 [29.0] -3.0 [10.0] -28.0 [12.0]	**-46.4, 38.1** **-46.8, 47.2** **-35.1, 27.7** **-45.3, 21.3** **-61.5, -14.5** **-32.8, 80.8** **-22.6, 16.6** **-51.5, 43.9**	**35.0** **32.3** **45.7** **35.5** **1.0** **19.4** **66.1** **6.6**	**63.6** **59.5** **77.7** **65.1** **6.7** **38.1** **94.5** **25.2**	So
Australia 2009 *BT version NR*	475 71 316 88	All <10kg 10-25kg >25kg	-5.5 [12.9]	-31.2, 20.4	62.0 68.0 63.0 53.0	90.0 88.0 93.0 81.0	Stewart
Australia 2010 *BT version NR*	1235 174 520 541	All <1yr 1-4yrs 5-10yrs	**-6.3 [14.6]** -7.2 [15.0] -7.0 **[11.7]** -5.3 **[16.7]**	**-34.8, 22.3** **-36, 22** **-30, 16** **-38, 58**	**46.8** **44.8** **52.8** **43.1**	**88.1** 83.0 91.0 87.0	Casey
Hong Kong 2010 BT1998	1199 449 471 279	All 1-5yrs 6-8yrs 9-11yrs	-1.9 **[14.1]** -1.3 **[13.2]** -1.9 **[14.2]** -2.9 **[15.3]**	-29.5, 25.7 -27.2, 24.7 -29.7, 25.9 -32.8, 27.1	58.4 65.3 56.7 56.3	89.0 91.5 90.1 83.0	Cattermole
Croatia 2010 *BT version NR*	151	All			58.0		Lulic
Canada 2011	243	All	-11.9 **[14.9]**	-41.1, 17.3	48.1	81.5	Bourdeau

BT version NR							
South Africa 2011 *BT version NR*	2832 1710 1122	All 1-5yrs 6-10yrs	-0.9 [14.3] -2.4 [13.8] 1.6 [14.7]	**-28.9, 27.1** **-29.4, 24.6** **-27.2, 30.4**	64.2 63.5 65.4	**83.7** **84.7** **82.4**	Geduld
USA 2011 BT2007B	657	All			**29.9**	46.9	Knight
USA 2011 BT2007B	372	All	-3.0 [15.3]	**-33.0, 27.0**	63.0	**80.0**	Rosenberg
USA 2012 *BT version NR*	1214	All	-6.1 [12.3]	**-30.1, 18.0**	37.4	53.9	Abdel-Rahman
Singapore 2012 BT2007B	875	All	-0.6 **[14.0]**	-28.1, 26.8	58.9	**72.3**	Loo
Canada 2012 BT2002A	6361	All	-7.1 [16.9]	-40.2, 26.0	56.3	84.5	Milne
Korea 2012 *BT version NR*	124095 19854 40612 63629	All Infant Preschool School	-4.8 [10.9] -4.9 [12.2] -4.0 [9.0] -5.5 [11.7]	**-26.1, 16.5** **-28.8, 19.0** **-21.6, 13.6** **-28.4, 17.4**	**59.6** **55.1** **68.8** **55.7**	**90.7** **87.1** **95.8** **87.8**	Park
USA 2012 *BT version NR*	118	All	-8.4 **[8.3]**	**-24.7, 7.9**	50.0	**91.9**	Sinha
Thailand 2012 *BT version NR*	595 111 372 104 8	All <10kg 10-25kg 25-40kg >40kg			56.1 60.4 59.7 43.3 0.0		Trakulsrichai
Botswana 2012 *BT version NR*	758	All			**35.0**	55.0	Wozniak
USA 2013 *BT version NR*	415	All	-4.1 [11.9]	**-27.4, 19.2**	58.6	90.8	Abdel-Rahman
USA 2013 *BT version NR*	976	All			53.5	78.0	Abdel-Rahman
Iran 2013 *BT version NR*	403	All			72.5		Akabarian
Sudan 2013 *BT version NR*	583 227 141 215	All Well nourish Malnourish Sev malnour			**15.1** 28.2 8.5 6.7		Clark
Australia 2013 BT2007B	24426 294 4421 14028 5683	All <1yr 1-5yrs 6-10yrs 11-14yrs			49.4 47.3 55.7 49.7 43.8		Graves
Kenya 2013 BT2007B	967	BT 2007B	2.2 **[12.7]**	**-22.7, 27.1**	**56.2**	**87.9**	House
Korea 2013	105072 104998 102316	BT 2005 A BT 2007 B BT 2011 A	-4.5 -3.9 -0.5		64.8 65.6 67.7		Suh
South Africa 2013 BT2007B	453 120 193 140	All <12kg 12-20kg >20kg	-3.8 [11.2] -2.4 [10.5] -3.0 [10.2] -6.0 [12.6]	**-25.8, 18.2** **-23.0, 18.2** **-23.0, 17.0** **-30.7, 18.7**	63.6 65.8 65.8 58.6	91.6 93.3 93.8 87.1	Wells
USA 2013 *BT version NR*	207	All			56.0	81.0	Young
Mali 2014 *BT version NR*	365	All	8.2 [10.4]	**-12.2, 28.6**	41.2	67.4	Dicko
India 2014 *BT version NR*	321	All	10.8 [16.3]	**-21.1, 42.7**	28.0	60.0	Batmanabane
Thailand 2014 BT2007A	3869	All	-3.6 [12.9]	**-28.9, 12.9**	58.0		Chiengkriwate
Australia 2014 BT2007B	2102 363 1739	All <1yr 1-5yrs	-0.4 [11.0] -5.9 [10.8] 0.6 [9.3]	**-25.7, 24.9** **-27.5, 15.7** **-20.6, 21.8**	**63.6** **57.7** **71.7**		Allison
USA 2015 BT2011A	400	All	-9.7 [19.6]	**-48.1, 28.7**	40.0	68.0	Abdel-Rahman
New Zealand 2015 BT2011A	305	All			73.4		Britnell
USA 2015 *BT version NR*	324 65 163 96	All <1yr 1-5yrs 6-12yrs	13.0 18.7 10.0 14.8		52.5	81.4	Chavez
India 2015 BT2007B	769	All	8.5 [14.6]	**-20.1, 37.1**	**43.8**		Asskaryar
India 2015	416	All	**0.3 [13.8]**	**-26.6, 27.3**	**53.1**		Asskaryar
Mexico 2015 *BT version NR*	815 236 343 236	All <10kg 10-18kg >18kg	-0.9 [15.9] 0.5 [20.8] -2.6 [12.1] -0.0 [15.1]	**32.1, 30.3** **-40.3, 41.3** **-26.3, 21.1** **-29.6, 29.6**	55.3 46.2 64.1 51.5		Khouli
Nigeria 2016 *BT version NR*	300	All	3.1 [19.5]	**-35.2, 41.3**	**46.3**		Aliyu
Zambia 2016 *BT version NR*	1282	All	4.3		47.3	75.1	Bowen
South Africa 2016 BT2011A	300 73 128 99	All <12kg 12-20kg >20kg	5.5 [13.4] 7.3 [12.5] 6.3 [11.2] 3.3 [16.3]	**-20.8, 31.8** **-17.2, 31.8** **-15.7, 28.3** **-28.6, 35.2**	54.0 60.3 53.1 50.5	84.3 86.3 87.5 78.8	Georgoulas

Korea 2016 BT2011A	906	All	6.4 **[18.0]**	**-31.1, 39.5**	60.6		Jung
USA 2016 BT2011A	2434	All			54.4		Lowe
Japan 2016 BT2007B	237	All			75.9	97.5	Nosaka
Australia 2016 *BT version NR*	160 18 78 60 5	All <1yr 1-5yrs 6-10yrs 11-14yrs			68.2 50.0 73.1 54.4 91.4		O'Leary
Multinational 2016 BT2007B BT2011A	453990	All	4.6 [10.4]	**-15.8, 25.0**	62.9	95.6*	Ralston
Malaysia 2016 *BT version NR*	1163	All	2.6 [20.7]	**-37.9, 43.2**	**36.8**		Sahar
South Africa 2017 BT2011A	328 30 90 212	All <12kg 12-20kg >20kg	-5.1 [16.2] 4.8 [8.9] -0.7 [12.3] -8.4 [17.4]	-36.9, 26.6 -12.7, 22.3 -24.8, 23.5 -42.5, 25.6	49.4 73.3 60.0 45.8	79.9 93.3 91.1 73.6	Wells

6.3 Epilogue

6.3.1 Major findings

This was the first study to have critically evaluated the functioning of the Broselow tape both as a weight estimation tool and a drug dosing guide. It was also the first meta-analysis focused on the Broselow tape specifically. The Broselow tape is generally considered to be the “gold standard” for weight estimation in children, even though it has been widely criticised for being inaccurate in both the high-income and low- and middle-income countries. This systematic review was therefore useful to evaluate the evidence supporting the ongoing use of the Broselow tape.

There were three major findings. The first was that the Broselow tape was more accurate than age formulas, but the Broselow tape was not as accurate as the Mercy method and the PAWPER tape. The previous chapters grouped weight estimation systems into three classes: the least accurate methods include healthcare provider guesses and age formulas; the moderate accuracy methods include length-only methods, such as the Broselow tape and the methods with the highest accuracy were the Mercy method, the PAWPER tape and parental estimates.

The second major finding was that there was no evidence to support the Broselow tape’s role as a drug dosing guide. The tape has never been tested without the use of additional reference material to provide supplementary comprehensive dosing information. When the tape was tested against other methods that included comprehensive information, the tape always proved inferior in drug dosing accuracy. This raises major doubts about the value of the tape in terms of objective evidence.

The third major finding was the importance of user training to ensure the correct use of the tape. Although the tape seems simple to use, many studies showed a high proportion of errors when utilised by undertrained healthcare providers. This was useful to provide a warning that all weight estimation systems are susceptible to human factor errors. Adequate training is thus essential for all weight estimation

systems. This will be covered in more detail in the last of the studies in this thesis, in Chapter 10.

6.3.2 This study in the context of the thesis

The previous two chapters showed that two of the most commonly-used weight estimation systems were not as accurate as the newer methods of weight estimation. This chapter also showed that there was less value in the drug dosing information that is printed on the Broselow tape than was generally assumed. The evidence from this systematic review was strongly indicative of the need for resuscitation aids which completely eliminate the need for any form of calculations. The Broselow tape did not satisfy these requirements. In the context of this research, this was very helpful as it showed that having this information on a tape is impossible because of size constraints for the tape as noted by Broselow himself [13]. This led to the conclusion that the tape should be used for weight estimation alone and a separate resuscitation aid should be used to provide comprehensive drug dosing information. The combination of these two elements – an accurate weight estimation tape and a dedicated companion resuscitation aid – would provide the optimum system to limit the risk of patient harm from medication errors.

One of the studies that was done as part of this research, but not included in the thesis, evaluated the accuracy of various models of habitus-modified Broselow tape weight estimation [14]. This showed that the Broselow tape could still have a role to play in the future of emergency weight estimation and drug dosing, but its methodology and content need to be updated based on the currently available evidence.

6.3.3 Link to the next article

Given the findings of the previous three chapters, it is clear that one-dimensional age-based or length-based weight estimation methods were substantially less accurate than the dual length- and habitus-based systems. This provided the

evidence to confirm the value of developing, refining and further researching the PAWPER system.

The next four chapters contain individual studies which were designed to identify and validate methods to enhance the functioning of the PAWPER system. The impetus for these studies were anecdotal clinical experiences with the tape as well as three international studies which provided useful insight into the vulnerabilities of the PAWPER system [15-17]. The first major challenge identified from using the PAWPER tape in other, older and more obese populations was the difficulty to provide weight estimations in taller children and severely obese children. The solution to this problem will be explored in Chapter 7 with the development of the PAWPER XL tape (extra-large and extra-long). The second challenge was the difference in the observers' perception of habitus [18]. This exposed the subjectivity of habitus assessment of the original tape's methodology. For this reason, more objective methods of habitus assessment were identified and explored in two further studies in Chapters 8 and 9.

These studies yielded useful information, which was very helpful in the development of enhanced methodologies for the PAWPER system.

6.3.4 References

[1] Schmidt AR, Haas T, Buehler KP, Ruetzler K, Weiss M, A S. A comparison of Broselow tape and pediatric emergency ruler. Accuracy of weight estimation. Notfall Rettungsmedizin. 2015;19:129-135.

[2] Young KD, Korotzer NC. Weight estimation methods in children: a systematic review. Annals of Emergency Medicine. 2016;68(4):441-451 e10.

[3] Stevens AD, Hernandez C, Jones S, Moreira ME, Blumen JR, Hopkins E, et al. Color-coded prefilled medication syringes decrease time to delivery and dosing errors in simulated prehospital pediatric resuscitations: a randomized crossover trial. Resuscitation. 2015;96:85-91.

[4] Both CP, Schmitz A, Buehler PK, Weiss M, Schmidt AR. How accurate are pediatric emergency tapes? A comparison of 4 emergency tapes with different length-based weight categorization. Pediatric Emergency Care. 2017.

[5] Moreira ME, Hernandez C, Stevens AD, Jones S, Sande M, Blumen JR, et al. Color-coded prefilled medication syringes decrease time to delivery and dosing error in simulated Emergency Department pediatric resuscitations. Annals of Emergency Medicine. 2015;66(2):97-106 e3.

[6] Benjamin L, Frush K, Shaw K, Shook JE, Snow SK, American Academy Of Pediatrics CoPEM, et al. Pediatric Medication Safety in the Emergency Department. Ann Emerg Med. 2018;71(3):e17-e24.

[7] Morgan N, Luo X, Fortner C, Frush K. Opportunities for performance improvement in relation to medication administration during paediatric stabilisation. Quality & Safety in Health Care. 2006;15:179-183.

[8] Hohenhaus SM, Frush KS. Pediatric patient safety in resuscitation simulations: lessons learned from studying clinician behavior. Pediatric Emergency Care. 2005;21(1):71-75.

[9] Clark MC, Lewis RJ, Fleischman RJ, Ogunniyi AA, Patel DS, Donaldson RI. Accuracy of the Broselow Tape in South Sudan, "The Hungriest Place on Earth". Academic Emergency Medicine. 2016;23(1):21-28.

[10] Asskaryar F, Shankar R. An Indian pediatric emergency weight estimation tool: prospective adjustment of the Broselow tape. International Journal of Emergency Medicine. 2015;8(1):78.

[11] Milne WK, Yasin A, Lubell R. Canadian children have outgrown the Broselow tape. Canadian Journal of Emergency Medicine. 2009;11:274.

[12] Nieman CT, Manacci CF, Super DM, Mancuso C, Fallon WF, Jr. Use of the Broselow tape may result in the underresuscitation of children. Academic Emergency Medicine. 2006;13(10):1011-9.

[13] Broselow J. From humble beginnings: the birth of the Broselow Tape 2012 [Accessed March 2018]. Available from: http://epmonthly.com/article/from-humble-beginnings-the-birth-of-the-broselow-tape/

[14] Wells M, Goldstein L, Bentley A. Accuracy of weight estimation by the Broselow tape is substantially improved by including a visual assessment of body habitus. Pediatric Research. 2018;83(1):83-92.

[15] O'Leary F, John-Denny B, McGarvey K, Hann A, Pegiazoglou I, Peat J. Estimating the weight of ethnically diverse children attending an Australian emergency department: a prospective, blinded, comparison of age-based and length-based tools including Mercy, PAWPER and Broselow. Archives of Disease in Childhood. 2017;102(1):46-52.

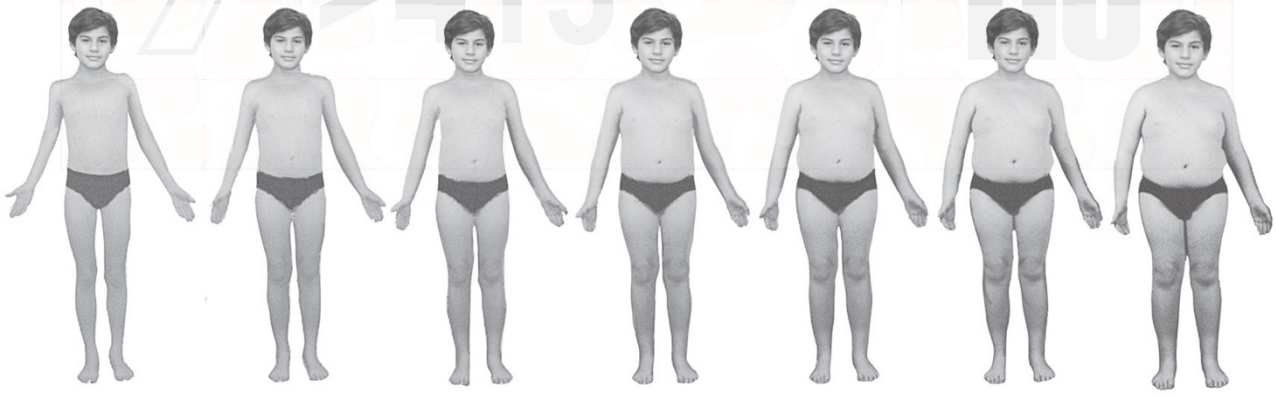
[16] Garcia CM, Meltzer JA, Chan KN, Cunningham SJ. A validation study of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape-a new weight estimation tool. Journal of Pediatrics. 2015;167(1):173-177.e1.

[17] Chavez H, Peterson R, Lo K, Arel M. Weight estimation in an inner city Pediatric Emergency Department: the effect of obesity. American Journal of Emergency Medicine. 2015;33(10):1364-1367.

[18] Goldstein L, Wells M. Fat is the new normal. Journal of Pediatrics. 2016;168:257.

Section 3

Refining the function of the PAWPER tape system



Chapter 7

A validation study of the PAWPER XL tape.

7.1 Prologue

7.1.1 Background

The previous chapters have provided the background information on the relative accuracy and characteristics of the various weight estimation systems. These studies showed that the PAWPER tape was one of the three most accurate weight estimation systems in the world, along with the Mercy method and parental estimates of weight [1]. This was a very encouraging finding. However, there were still some subgroups in which the performance of the tape did not achieve the stringent acceptable outcome criteria that was desired. Since one of the major purposes in this thesis was to develop and optimise the PAWPER systems, these limitations which had been identified needed to be addressed.

Studies on the original PAWPER tape in Australia and the USA revealed four important facts [2-4].

- The tape was not as accurate in populations with a high prevalence of obesity (>30% of children) as in those populations with a lower prevalence of obesity (<20% of children).
- The tape could not provide a weight estimation for children taller (or longer) than 153 cm. This meant that many children around the age of 12 were weight estimation “orphans”. While most children presenting to general paediatric Emergency Departments are below the age of 5 years, many departments, and especially trauma centres, see a large number of older children [5]. One quarter of all Emergency Department visits in the United Kingdom are by children under the age of 19 and about 40% of these visits are by children over the age of 10 years [6]. Given this significant proportion of children over 10 years of age, it was appropriate to design a weight estimation system that would be usable in all children.
- The assignment of a habitus score, which is an important determinant of accuracy of the PAWPER tape, was inconsistent and inaccurate in populations with a high prevalence of overweight and obese children. This is

because the visual “gestalt” method requires the user to accurately identify normal weight children and adjust the habitus score up or down for overweight and underweight children, respectively. This is more difficult in populations where “fat is the new normal” [7, 8]. This theme will be covered further in Chapter 8.

- Finally, despite the abovementioned limitations, the PAWPER tape still performed better than any other weight estimation system in these studies.

An understanding of these limitations assisted in the development of the PAWPER XL tape. The new concepts needed to be tested, however.

The PAWPER tape had previously been shown to be accurate in underweight and normal weight children, but with the focus of the new tape on obese children, the matter of the correct drug dosing strategies for these children had to be considered. While some drugs should still be dosed to TBW in obese children, others are better dosed to IBW [9]. An estimate of IBW would therefore also be useful. Since IBW is very closely linked to length in children, it would not be difficult to use a length-based tape for this purpose. The hypothesis that the Habitus Score 3 weight for a child would accurately predict their IBW needed to be tested.

7.1.2 Purpose of the study

This study was chiefly intended to provide a published description of the functioning of the PAWPER XL tape and an initial and preliminary assessment of its performance. The primary purpose of this study was to evaluate the performance of the PAWPER XL tape in a sample of children that included children “too tall” for the original PAWPER tape and children who were obese or severely obese. These categories of children have previously been identified as being at risk of suboptimal weight estimation by the original PAWPER tape. The study was also planned to directly compare the new PAWPER XL tape with the original PAWPER tape to evaluate its effectiveness in the overall sample as well as in subgroups of overweight and obese children.

The secondary purpose of this study was to see whether the PAWPER XL tape could accurately estimate IBW in obese children. This is would be useful because, in obese children, IBW is required for dose calculations of hydrophilic drugs and TBW is required for lipophilic drugs [10].

This study addressed Objective 3 (to develop a new version of or methodology for the PAWPER tape based on the findings of previous studies; to compare the weight estimation performance of this new version against other weight estimation systems; and to establish whether it could accurately predict both total and ideal body weight in obese children).

7.2 Published paper

Wells M, Goldstein L, Bentley A. A validation study of the PAWPER XL tape: accurate estimation of both total and ideal body weight in children up to 16 years of age. Trauma and Emergency Care. 2017;2(4):1-8.

The PDF of the published article is included.

Trauma and Emergency Care

Research Article

A validation study of the PAWPER XL tape: accurate estimation of both total and ideal body weight in children up to 16 years of age

Mike Wells[1]*, Lara Goldstein[1] and Alison Bentley[1]

[1]Division of Emergency Medicine,Faculty of Health Sciences, University of the Witwatersrand, South Africa

Abstract

Background: The PAWPER tape has proved to be one of the most accurate weight estimation systems available, but its reduced accuracy in obese children and relative shortness (153cm) limit its functioning. The PAWPER tape was redeveloped as the PAWPER XL tape, to provide additional capacity for estimating weight in obese children and taller children (for extra-length and extra-large children). The aim of this study was to evaluate the accuracy of the PAWPER XL tape in estimating total body weight (TBW) and ideal body weight (IBW) in a population with a high prevalence of underweight and obese children.

Methods: Estimations of TBW and IBW were obtained using the Broselow tape, the Mercy method, the original PAWPER tape and the new PAWPER XL tape in a convenience sample of 332 Emergency Department children. These predicted weights were compared to actual weight and calculated IBW.

Results: The percentage of TBW estimates within 10% of actual weight (PW10) for the PAWPER XL tape, the PAWPER tape, the Mercy method and the Broselow tape was 83.4%, 81.8%, 63.9% and 57.1% respectively. For IBW the PW10 for the PAWPER XL tape, the PAWPER tape and the Broselow tape was 87.9%, 86.7% and 80.0% respectively.

Conclusions: The PAWPER XL tape estimated both TBW and IBW extremely accurately, significantly better than the other weight estimation systems. The increased length and number of habitus score categories of the PAWPER XL tape enabled it to outperform the PAWPER tape in children >153cm in length and in severely obese children.

Abbreviations: Total body weight-TBW; Ideal body weight-IBW; Emergency Department-ED

Introduction

Obtaining an estimate of weight forms part of the emergency management of critically ill or injured children, as most drug doses, and many other interventions, are based on weight. The effectiveness and safety of treatment is ultimately dependent on the accuracy of weight estimation, as enough drug must be administered to ensure efficacy but not too much to cause toxic or unwanted side effects [1]. Although some drugs (such as opioids) can be titrated to effect, some drugs (such as anticonvulsants) require an accurate first-time administration to ensure optimum outcomes [2]. The most accurate weight-estimation systems available should therefore be used, and the continued use of weight estimation methods that are known to be inaccurate should be considered to be poor medical practice [3].

The need for a modification of the original PAWPER system became apparent from the results of local and international research [4-6]. While the PAWPER system still proved to be more accurate than other weight estimation systems in these studies, some limitations of the tape were identified. Firstly, the original PAWPER tape was too short at 153cm to provide weight estimations for some children. Studies on children "too tall for the tape" (referring primarily to the Broselow tape) showed that these children should not be treated as adults in terms of drug dosing [7]. Children as young as 9 or 10 years of age could fall beyond the length of the Broselow tape (145cm) and therefore put at risk of significant overdosing errors if dosed as adults, or underdoing if dosed according to the last Broselow weight. Although the original PAWPER tape is slightly longer than the Broselow tape, it still failed to provide weight estimations for some older children. We decided that it was essential that the PAWPER system be adapted to provide an accurate weight estimation in children up to the age of at least 16 years. Secondly, the original PAWPER system failed to provide accurate weight estimations for children who were severely obese. Two studies from populations in the USA with a high prevalence of obesity showed that the PAWPER tape underestimated weight in obese children [4,5]. This could lead to dosing errors in these children.

In order to avoid underestimation of weight in taller, obese and severely obese children, the PAWPER XL tape was developed for "extra-length" and "extra-large" children (see Figure 1). The major changes between the original PAWPER tape system and the new PAWPER XL tape system were: the length of the tape was increased from 153cm to 180cm; the number of habitus score categories was increased from five to seven, with HS6 and HS7 added for severely obese children (children on the 97th and 99th centiles of weight-for-length); changes were made

Correspondence to: Mike Wells, Division of Emergency Medicine, Faculty of Health Sciences, University of the Witwatersrand, 5 York Road, Parktown Johannesburg, South Africa, Tel: +27824910369; E-mail: mike.wells@emergencymedicine.co.za

Key words: *PAWPER XL tape, medical emergencies, emergency department, body weight, fat-free mass*

Received: May 15, 2017; **Accepted:** June 07, 2017; **Published:** June 09, 2017



Figure 1. The PAWPER XL tape. This tape is a 1800mmx40mm tape with 13 colour-zones and 34 weight divisions, from 3kg to 70kg. The lower panel shows the appearance of the 60kg and 65kg ideal body weight segments, which fall into the zebra-stripe and giraffe-spot colour-zones. The seven habitus score categories along with their predicted weights are shown. The ideal body weight is predicted by the habitus score 3 weights.

to HS1 and HS2 weights in shorter children and HS 4 and 5 weights in taller children to improve accuracy of weight estimations; and each weight segment of the PAWPER XL tape was labelled with the ideal body weight (IBW), which corresponded to the HS3 weight at that length. How the tape is used is shown in Figure 2.

With the increasing worldwide prevalence of obesity in children total body weight (TBW) may not be the best weight-descriptor for dose calculations for all drugs in all children [8]. The problem of potential overdosing of obese children poses challenges during dose calculations based on TBW for drugs with serious side-effects, such as analgesic, anti-convulsant or anti-arrhythmic drugs. For these drugs, a dosing scalar such as ideal body weight (IBW) is recommended instead of TBW during weight-based dose calculations [9,10]. The use of IBW as a weight-descriptor is usually most applicable in obese children when dosing hydrophilic medications with a small volume of distribution (e.g. adrenaline).

The aim of this study was to establish the accuracy of the PAWPER XL tape in estimating both TBW and IBW in a sample of children with a high prevalence of both underweight and obesity. The secondary objective was to compare the accuracy of the PAWPER XL tape with the original PAWPER tape, the Broselow tape and the Mercy method.

Methods

Study design and setting

This was a prospective, cross-sectional study conducted in the Emergency Department (ED) of a medium-sized, academically aligned private hospital in Johannesburg, South Africa, which provides service to approximately 6000 paediatric patients per year. Ethics approval was obtained from the institutional review board as well as from the Human Research Ethics Committee of the University of the Witwatersrand.

Selection of participants

A convenience sample of 332 children from one month to 16 years of age who presented to the ED or who were an in-patient at the study hospital between October 2014 and January 2015 were enrolled. Only children whose participation did not interfere with emergent medical treatment were eligible for inclusion. Written informed consent and assent (where applicable) was obtained for all participants.

Methods and measurements

Data was collected by one of the researchers (MW or LG) with the same procedure followed for each child. Basic demographic data was obtained and the children then changed into a hospital gown for the subsequent measurements. The child was positioned supine on the examination bed for weight estimation by the Broselow and original PAWPER tapes according to the directions on the tapes. Weight estimations for the PAWPER XL tape were generated from measurements of length and a visual assessment of body habitus. Basic anthropometric measurements were also obtained with the child in a supine position: length, mid-arm circumference (MAC) and humerus length. The child was finally weighed on a calibrated digital scale (Tanita SC-240 Body Composition Analyser).

The anthropometric measurements were used to calculate a number of additional measures. An estimation of TBW was obtained

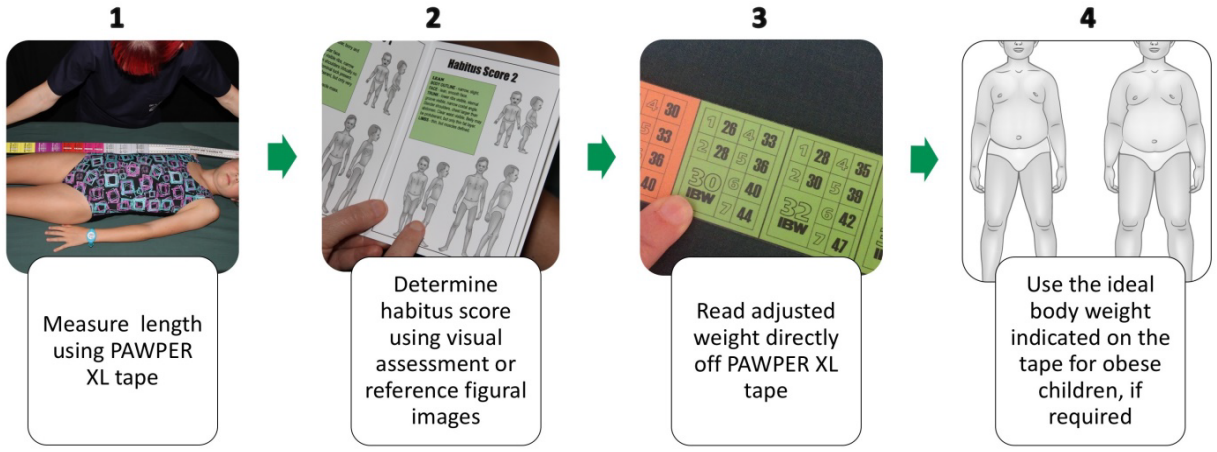


Figure 2. How to use the PAWPER XL tape. The steps required to use the PAWPER XL tape are shown. Both total body weight and ideal body weight can be read directly off the tape.

from the MAC and humerus-length data using the Mercy method [11]. Body mass index (BMI) and the corresponding BMI-for-age Z-scores were determined. An estimate of IBW was determined using the BMI50 method [12]. IBW estimation by the Broselow tape was considered to be the same as the TBW estimation; IBW estimation by the PAWPER XL tape was obtained using the HS3 weight at the child's length.

Analysis

Each weight-estimation system was compared with measured TBW and calculated IBW using both parametric and non-parametric statistical methods. Three primary statistical measures were used to assess performance: mean percentage error (MPE) was calculated to quantify the overall estimation bias; the 95% limits of agreement of the mean percentage error represented the estimation precision; and the percentage of weight estimations that fell within 10% (PW10) and 20% (PW20) of the measured weight-descriptor represented overall accuracy.

Subgroup analyses were performed in three weight categories (<10kg, 10 to 25kg and >25kg) and three habitus categories based on BMI-for-age Z-scores ("thin" children Z-score <-2.0, "overweight" children Z-score >+2.0 and "average" children in between). Subgroup analyses were also performed in children with a length >145cm and HS≥5.

The data was also graphically represented using a modified Bland-Altman plot. Comparisons of accuracy were made using the McNemar or Fisher exact tests. A p <0.05 was considered to be significant for all statistical tests.

Outcome measures

The primary outcome measure was the performance of each of the weight-estimation models when compared to actual weight. A PW10 of >70% and a PW20>95% was regarded as an acceptable level of accuracy. This benchmark was derived from previously suggested criteria, as well as the level of accuracy commonly achieved by newer dual length- and habitus-based weight estimation systems [13]. This benchmark was also considered to indicate an acceptable accuracy of IBW estimation.

Results

Characteristics of study participants

A total of 332 children were included in the study. The basic demographic information is shown in Table 1. The BMI-for-age distribution showed a population with a sizeable number of underweight children (15.3%) and overweight or obese children (32.5%) in which to test the weight estimation systems. There was also a high prevalence of children (20.5%) whose TBW and IBW differed by more than 20%.

Main results

The performances of each of the weight-estimation systems against TBW and IBW are shown in Table 2. With regards to estimating TBW, the Broselow tape and the Mercy method achieved an intermediate degree of accuracy (PW10s of 57.1 and 63.9 respectively) while the PAWPER tape and PAWPER XL tape achieved an extremely high degree of accuracy (PW10 81.8% and 83.4% respectively). The modified Bland & Altman percentage error plots for each of the weight estimation systems are shown in Figure 3.

There were 43 children (13.0% of the sample) >145cm in length who were too tall for the Broselow tape and 24 children (7.2% of the sample) >153cm in length who were too tall for the PAWPER tape. Only the Mercy method and the PAWPER XL tape could provide weight estimations in these children, which were similar in accuracy to that in shorter children (PW10s of 70.3% and 83.3% respectively, Fisher exact test, p<0.0001).

There were 34 obese or severely obese children (10.2% of the sample) in whom the weight estimation by the original PAWPER system was handicapped because of the maximum habitus score limitation. The Mercy method and the PAWPER XL tape again were the best performers in these children, but with only a moderate degree of accuracy (PW10s of 53.3% each) while the Broselow tape and the original PAWPER tape had very low accuracy (PW10s of 0% and 16.7% respectively).

While the Broselow tape was substantially more accurate in average weight children than in underweight, overweight and obese children, the Mercy method, the PAWPER tape and the PAWPER XL tape had a more consistent performance across the spectrum of body-types and weight categories. The relatively poorer performance of all the systems in children <10kg was partly because of the high incidence of obesity and severe obesity in this group.

The Broselow tape, the original PAWPER tape and the PAWPER XL tape all predicted IBW extremely well in obese children (PW10 80.0, 86.7% and 87.9% respectively, Fisher exact test, p<0.0001 for comparison between Broselow tape and PAWPER tapes).

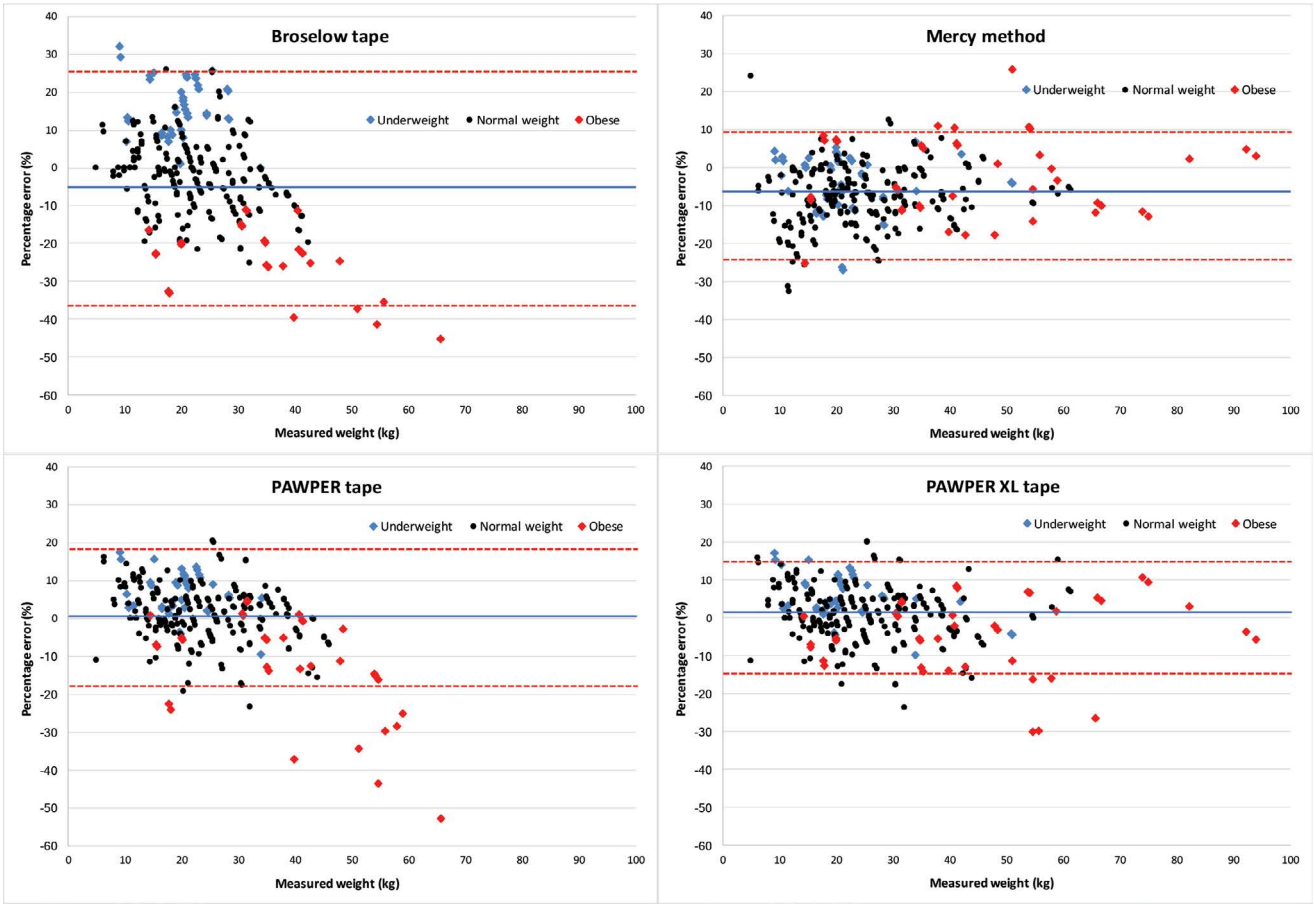


Figure 3. Modified Bland & Altman plots for each weight estimation system. A percentage error method was used rather than actual differences, because of the wide range of weights of participating children. The solid blue lines indicate the mean percentage error (MPE) while the dashed red lines indicate the 95% limits of agreement. Underweight children, normal weight and obese children are shown with different markers on the plot.

Table 1. Description of the study population: demographic information with body composition data. Subgroup data are provided using a body mass index (BMI) classification. Since body composition reference data have not been well established in children, and especially in younger children, most analyses in this study made use of pragmatic limits that might affect drug dosing decisions: children were considered to be significantly obese when their total body weight (TBW) was >120% of ideal body weight (IBW) (which roughly corresponds to a Z-score of +2) as this would require the use of IBW as a drug dosing descriptor for some drugs. Likewise, children were considered significantly underweight when their TBW was <80% IBW (which approximates a Z-score of -2) as the use of IBW would result in a critical overdosing error. Medians and interquartile ranges are shown.

	All	Underweight	Thin	Normal	Overweight	Obese	Severely obese
Number	332	18 (5.4)	33 (9.9)	173 (52.1)	74 (22.3)	20 (6.0)	14 (4.2)
Sex (male)	154 (46.4)	4 (22.2)	15 (45.5)	93 (53.7)	25 (33.8)	12 (60.0)	5 (35.7)
Age (years)	7.2 (4.5, 9.3)	6.4 (3.5, 8.0)	9.1 (5.7, 13.1)	7.3 (4.9, 10.1)	7.8 (5.0, 9.4)	6.7 (4.9, 10.0)	3.0 (1.7, 6.9)
BMI (kg/m²)	16.7 (15.2, 18.8)	13.2 (12.9, 13.4)	16.3 (15.6, 18.6)	16.2 (15.4, 17.3)	18.6 (17.8, 20.0)	19.9 (19.3, 26.6)	22.6 (20.5, 26.4)
BMI-for-age Z-score	0.4 (-0.5, 1.1)	-2.7 (-3.5, -2.2)	0.0 (0.0, 0.0)	0.3 (-0.3, 0.6)	1.4 (1.1, 1.6)	2.1 (2.0, 2.2)	2.9 (2.8, 3.5)
IBW (kg)	23.2 (17.7, 30.7)	22.2 (14.9, 23.5)	26.4 (21.2, 43.0)	22.8 (18.3, 31.5)	25.9 (18.9, 31.5)	23.7 (18.2, 35.9)	14.2 (11.3, 35.9)
IBW >120% TBW n (%)	10 (3.0)	9 (50.0)	1 (3.0)	-	-	-	-
TBW>120% IBW n (%)	58 (17.5)	-	-	1 (0.6)	25 (33.8)	18 (90.0)	14 (100)
HS	3 (3, 4)	2 (2, 2)	3 (3, 3.75)	3 (3, 4)	4 (4, 4.8)	4.5 (4, 5)	5 (5, 6)
TBW (kg)	23.4 (17.6, 33.8)	18.3 (12.3, 20.0)	26.6 (21.1, 43.1)	23.4 (18.4, 31.4)	30.6 (21.6, 38.8)	30.8 (21.3, 56.3)	19.0 (15.5, 39.3)

The results of the statistical comparisons between the weight estimation systems are shown in Table 3. The Broselow tape was statistically inferior to the other systems in virtually every analysis. The Mercy method was statistically inferior to the PAWPER tape and PAWPER XL tape in the overall sample, in children >10kg of weight and in children of average habitus, but was significantly superior to the PAWPER tape in obese children. The PAWPER XL tape outperformed the PAWPER tape in the subgroup of overweight and obese children.

Discussion

The performance of the weight estimation systems – TBW estimation

The PAWPER XL tape performed very well in this study and achieved the acceptable outcome criteria. It performed very well in every subgroup except for the group of severely obese children. Although there is no previous data from studies on the PAWPER

Table 2. Outcomes of the weight-estimation systems against total body weight (TBW) and ideal body weight (IBW). Subgroups by weight, thinness (IBW>120% TBW) and fatness (TBW > 120% IBW) are shown. The performance of estimations of IBW are only shown for obese children (in whom these descriptors might be used). To evaluate the performance of the PAWPER XL tape, the outcomes for each system are shown for comparison for children "too tall for the original tape" (>145cm) and those with habitus scores above the obese range (HS 6 and HS 7). A description of the terminology for outcomes is provided in the lowest panel.

Total body weight (TBW) estimation outcomes							
		N	**PW10**	**PW20**	**MPE**	**LLOA**	**ULOA**
Broselow tape	All	289	57.1	86.5	-1.9	-27.8	24.0
	<10kg	11	72.7	81.8	6.7	-18.1	31.4
	10-25kg	168	62.5	90.5	0.9	-21.9	23.7
	>25kg	110	47.3	80.9	-7.0	-34.2	20.2
	Thin	41	19.5	58.5	17.5	4.5	30.5
	Average	202	75.7	99.0	-1.5	-17.2	14.2
	Overweight	46	8.7	56.5	-20.9	-38.5	-3.2
	>145cm	43	-	-	-	-	-
	HS≥5	29	0	0	-36.4	-56.4	-16.4
Mercy method	All	332	63.9	94.3	-6.7	-23.1	9.6
	<10kg	11	54.5	90.9	-4.2	-27.4	19.0
	10-25kg	168	64.9	92.3	-7.5	-23.8	8.9
	>25kg	153	63.4	96.7	-6.1	-21.8	9.7
	Thin	46	78.3	93.5	-3.7	-19.9	12.5
	Average	218	62.3	94.7	-8.0	-22.6	6.6
	Overweight	58	58.6	93.1	-4.2	-24.6	16.2
	>145cm	43	70.8	100	-4.8	-18.4	8.7
	HS≥5	34	53.3	93.3	-1.1	-25.2	22.9
PAWPER tape	All	308	81.8	96.2	0.5	-17.9	18.9
	<10kg	11	54.5	100	8.7	-7.1	24.4
	10-25kg	168	85.7	98.8	1.9	-11.6	15.4
	>25kg	112	78.6	92.0	-2.4	-25.2	20.4
	Thin	41	58.5	97.6	9.2	-0.7	19.0
	Average	202	93.1	100	1.7	-9.4	12.7
	Overweight	48	54.2	79.2	-11.8	-36.7	13.1
	>153cm	24	-	-	-	-	-
	HS≥5	31	16.7	33.3	-24.6	-56.3	7.1
PAWPER XL tape	All	332	83.4	98.5	1.1	-13.9	16.1
	<10kg	11	54.5	100	8.7	-7.1	24.4
	10-25kg	168	86.3	100	2.0	-10.4	14.3
	>25kg	153	82.4	96.7	-0.4	-17.1	16.3
	Thin	46	70.4	97.8	7.8	-3.2	18.9
	Average	218	91.7	100	1.5	-10.0	13.1
	Overweight	58	70.8	93.1	-6.0	-24.5	12.5
	>145cm	43	83.3	97.9	-0.3	-16.8	16.2
	HS≥5	34	53.3	86.7	-6.2	-30.0	17.7
Ideal body weight (TBW) estimation outcomes							
		N	**PW10**	**PW20**	**MPE**	**LLOA**	**ULOA**
Broselow tape	Obese	30	80.0	100	-3.5	-26.2	19.3
PAWPER tape	Obese	30	86.7	100	4.5	-6.1	15.1
PAWPER XL tape	Obese	34	87.9	100	4.8	-4.5	14.1

Description of accuracy outcomes

PW10	**Descriptor**	**PW20**	**Descriptor**
<30%	Critically inaccurate	<80%	Very high critical error rate
30-40%	Very inaccurate	80-90%	High critical error rate
40-50%	Inaccurate		
50-60%	Moderately accurate	90-95%	Moderate critical error rate
60-70%	Accurate		
70-80%	Very accurate	>95%	Low critical error rate
80-90%	Extremely accurate		
>90%	Exceptionally accurate		

Table 3. Outcomes of the statistical tests between the weight estimation systems for the whole sample and for the indicated subgroups. The Fisher exact test was used. The abbreviations shown in the matrix indicate which system was more accurate in the paired tests. Abbreviations: BT – Broselow tape, MM – Mercy method; PT – original PAWPER tape, PTXL – PAWPER XL tape, NS – not significant.

All			
	Broselow tape	**Mercy method**	**PAWPER tape**
Mercy method	MM (p=0.0041)		
PAWPER tape	PT (p<0.0001)	PT (p<0.0001)	
PAWPER XL tape	PTXL (p<0.0001)	PTXL (p<0.0001)	NS
Weight <10kg			
	Broselow tape	**Mercy method**	**PAWPER tape**
Mercy method	NS		
PAWPER tape	NS	NS	
PAWPER XL tape	NS	NS	NS
Weight 10 to 25kg			
	Broselow tape	**Mercy method**	**PAWPER tape**
Mercy method	NS		
PAWPER tape	PT (p<0.0001)	PT (p<0.0001)	
PAWPER XL tape	PTXL (p<0.0001)	PTXL (p<0.0001)	NS
Weight >25kg			
	Broselow tape	**Mercy method**	**PAWPER tape**
Mercy method	MM (p<0.0001)		
PAWPER tape	PT (p<0.0001)	PT (p<0.0001)	
PAWPER XL tape	PTXL (p<0.0001)	PTXL (p<0.0001)	NS
Thin children			
	Broselow tape	**Mercy method**	**PAWPER tape**
Mercy method	MM (p<0.0001)		
PAWPER tape	PT (p<0.0001)	NS	
PAWPER XL tape	PTXL (p<0.0001)	NS	NS
Average children			
	Broselow tape	**Mercy method**	**PAWPER tape**
Mercy method	NS		
PAWPER tape	PT (p<0.0001)	PT (p<0.0001)	
PAWPER XL tape	PTXL (p<0.0001)	PTXL (p<0.0001)	NS
Overweight children			
	Broselow tape	**Mercy method**	**PAWPER tape**
Mercy method	MM (p<0.0001)		
PAWPER tape	PT (p<0.0001)	NS	
PAWPER XL tape	PTXL (p<0.0001)	NS	PTXL (p<0.0001)
Obese children			
	Broselow tape	**Mercy method**	**PAWPER tape**
Mercy method	MM (p<0.0001)		
PAWPER tape	PT (p<0.0001)	MM (p<0.0001)	
PAWPER XL tape	PTXL (p<0.0001)	NS	PTXL (p<0.0001)

XL tape, it performed better in obese children in this study than the original PAWPER tape did in obese populations in previous studies [4,5]. The use of figural reference images to increase the accuracy of habitus assessment may increase the performance of the PAWPER XL tape further.

The Broselow tape achieved moderate accuracy in this study but did not meet the acceptable outcome criteria. Although the Broselow tape accurately estimated weight in children of average habitus (those with a TBW similar to IBW), it was critically inaccurate in underweight and overweight children and failed to estimate weight within 20% of actual weight in every obese child in this study. This has been frequently reported previously, with potentially dangerous overestimation of weight in low- and middle-income countries and substantial underestimation of weight in high-income populations [14-18]. The accuracy was highest in the youngest children, falling off significantly in children >25kg. This is also a pattern repeatedly reported previously [19]. The Broselow tape was simply not accurate enough and could not provide a weight estimation in a substantial number of children who were too tall for the tape. Given its repeated failure to achieve acceptable accuracy in this study and many previous studies, the role of the Broselow tape as the gold standard of weight estimation in children needs to be reconsidered.

The Mercy method achieved a lower accuracy in this study than in many previous studies (PW10s generally above 70%) [11,20,21]. This is probably because children were measured supine in this study to simulate how the method would be used during emergency care. Differences in skill and experience in anthropometry may also have played a role, however, as has been previously noted in studies on the Mercy method [20,22]. Improving the performance of the Mercy method, when used by novices and ordinary clinicians, needs to be explored further. Its use during actual emergency care also needs to be evaluated.

The original PAWPER tape achieved an exceptionally high overall level of accuracy in this study sample, similar to that previously reported from non-obese populations [6,23,24]. The accuracy in obese children was poor, however, similar to previous findings from studies in populations with a high prevalence of obesity [4,5]. The possible explanations for poor accuracy are, probably, a failure of the users to visually assess habitus accurately, leading to an underestimation of a child's weight status; and an inherent inability of the tape to provide weight estimations for obese and severely obese children [25]. These shortcomings provided the motive to develop the PAWPER XL tape.

A comparison between the PAWPER XL tape and the other methods

The overall performance of the PAWPER systems was better than that of the Mercy method, which was better than that of the Broselow tape. These findings echo previous studies in which these three methods have been compared [5,6,23]. Given the consistently higher accuracy of the dual length- and habitus-based weight estimation systems (such as the Mercy method and the PAWPER tapes) when compared with length-only systems (such as the Broselow tape) it is hard to justify the continued use of the Broselow tape.

The PAWPER XL tape had the best performance of all the systems in every subgroup except the <10kg group. In this group, the Broselow tape performed best because it overestimated weight less than the other methods. The differences between the PAWPER XL tape and the PAWPER tape were small overall, but it was the areas of difference between the tapes that was of most interest: in the >145cm group and the obese and severely obese children. In taller children, even with a relatively small subgroup size, the PAWPER XL tape was extremely accurate and as accurate as in shorter children. The accuracy in the subgroup of obese children was significantly and substantially better than the Broselow tape and the PAWPER tape, but similar to the Mercy method. Since the PAWPER XL tape still underestimated the weight in obese children, it might be possible to recalibrate the system to further improve accuracy, but there was no clear indication of how the Mercy method could be improved.

Children "too tall for the tape" and "too fat for the tape"

One of the important measures of a weight estimation system is the restrictions to its use. The Broselow tape and the original PAWPER tape had a substantial number of "weight estimation orphans" for whom another method would have had to be used to estimate their weight in a clinical setting. The use of either adult weight for these children,

or the highest weight on the tape has been shown to be erroneous and potentially harmful [7]. The Mercy method and the PAWPER XL tape were able to provide weight estimations for all children in this study: both these systems are thus likely to work well in children up to approximately 16 years of age, making them useful beyond the restrictions of most other weight estimation systems.

The changes in the PAWPER XL tape, increased to a length of 180cm, showed excellent performance in the subgroup of taller children. With its habitus scores increased to seven from the original PAWPER tape's five, it showed reasonable performance in morbidly obese children but further assessment in a larger sample will still be required.

The performance of the weight estimation systems – IBW estimation

The Broselow tape, the original PAWPER tape and the PAWPER XL tape predicted IBW extremely accurately in obese children. Although it has previously been suggested that length-based methods could predict IBW, this is the first study to confirm that assumption [26]. The Mercy method, despite being accurate at predicting TBW, does not have a mechanism by which to estimate IBW: the use of humerus-length as a surrogate for recumbent length prevents easy, direct prediction of IBW.

The significance of IBW estimation

As the Pediatric Advanced Life Support (PALS) guidelines of the American Heart Association affirm, there is no clarity on whether drug doses must be adjusted during the resuscitation of obese children and there is no empirical evidence from which to create guidelines [27]. Their own guidelines are somewhat self-contradictory, however, as they regard the use of either actual body weight (TBW) or length-based weight estimations (IBW) for drug dosing as equivalent. It is clear from the findings of this study that TBW and IBW are not interchangeable: a fifth of the study population had more than a 20% difference between TBW and IBW. The use of IBW as a scalar will, therefore, result in medication overdoses in underweight children [28]. Similarly, the use of TBW in obese children will result in significant overdosing of hydrophilic medications. Although there is no clear evidence of harm from medication errors arising from the incorrect use of TBW or IBW as a dose scalar, there has been speculation that these errors may have led to poorer outcomes following cardiac arrest in obese children [29]. The theoretical considerations about the use of appropriate dose scalars in obese children must be given some credence, especially if both TBW and IBW can be rapidly and accurately estimated in an emergency [30].

Limitations

Although this was a preliminary study to evaluate the PAWPER XL tape, the subgroup sample sizes were quite small. Further research is still required in a sample with a greater number of obese children and taller children. Secondly, the process of assigning a habitus score using a visual assessment of habitus was subjective and may be a significant contributor to error, especially in obese children: this was not evaluated in this study. The use of more objective methods of habitus assessment, such as the use of figural reference images, need to be evaluated. Thirdly, the impact of inaccurate weight estimations on outcomes is not known, nor is the benefit of using IBW over TBW for appropriate drugs in obese children. This needs to be established to guide the stringency of further development of weight estimation systems.

Conclusions

The PAWPER XL tape was extremely accurate in this study and surpassed the acceptable accuracy benchmark. It was more accurate than the original PAWPER tape because of its expanded length and modified habitus score categories. Although the PAWPER XL tape was more accurate than the other systems in severely obese children, it did not achieve the acceptable outcome criteria in this subgroup. The Broselow tape performed poorly, as in multiple previous studies and was simply not accurate enough and not long enough. The continued use of the tape needs to be carefully reconsidered. The Mercy method was moderately accurate, but its accuracy appears to be dependent on patient positioning and the experience of its users. This also needs to be explored in future research. The Broselow tape, the PAWPER tape and the PAWPER XL tape predicted IBW with a high degree of accuracy. The significance and usefulness of an easily-determined IBW needs to be determined with some urgency, to ensure that emergency drug dosing in children is optimised and medication errors avoided.

References

1. Wells M, Goldstein L, Bentley A (2017) It is time to abandon age-based emergency weight estimation in children! A failed validation of 20 different age-based formulas. *Resuscitation* 116:73-83. [Crossref]
2. Manno EM (2011) Status epilepticus: current treatment strategies. *Neurohospitalist.* 1(1):23-31. [Crossref]
3. Luscombe M (2005) "Kid's aren't what they used to be": A study of paediatric patients' weights and their relationship to current weight estimation formulae. *British Journal of Anaesthesia* 95: 578.
4. Garcia CM, Meltzer JA, Chan KN, Cunningham SJ (2015) A validation study of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape-a new weight estimation tool. *Journal of Pediatrics*. 167: 173-177.e1. [Crossref]
5. Chavez H, Peterson R, Lo K, Arel M (2015) Weight estimation in an inner city Pediatric Emergency Department: the effect of obesity. *American Journal of Emergency Medicine*. 33: 1364-1367.
6. O'Leary F, John-Denny B, McGarvey K, Hann A, Pegiazoglou I, et al. (2017) Estimating the weight of ethnically diverse children attending an Australian emergency department: a prospective, blinded, comparison of age-based and length-based tools including Mercy, PAWPER and Broselow. *Archives of Disease in Childhood*. 102: 46-52.
7. Cattermole GN, Leung PY, Graham CA, Rainer TH (2014) Too tall for the tape: the weight of schoolchildren who do not fit the Broselow tape. *Emerg Med J* 31: 541-544. [Crossref]
8. Kendrick JG, Carr RR, Ensom MH (2010) Pharmacokinetics and drug dosing in obese children. *J Pediatr Pharmacol Ther* 15: 94-109. [Crossref]
9. Rowe S, Siegel D, Benjamin DK, Jr., Best Pharmaceuticals for Children Act - Pediatric Trials Network Administrative Core C (2015) Gaps in drug dosing for obese children: a systematic review of commonly prescribed emergency care medications. *Clinical Therapeutics* 37: 1924-1932. [Crossref]
10. Ross EL, Heizer J, Mixon MA, Jorgensen J, Valdez CA, et al. (2015) Development of recommendations for dosing of commonly prescribed medications in critically ill obese children. *American Journal of Health-System Pharmacy* 72: 542-556. [Crossref]
11. Abdel-Rahman SM, Ridge AL (2012) An improved pediatric weight estimation strategy. *Open Medical Devices Journal* 4: 87-97.
12. Phillips S, Edlbeck A, Kirby M, Goday P (2007) Ideal body weight in children. *Nutr Clin Pract* 22: 240-245. [Crossref]
13. Stewart D (2009) Accuracy of the Broselow tape for estimating paediatric weight in two Australian Emergency Departments: University of Sydney.
14. Batmanabane G, Kumar JP, Dikshit R, Abdel-Rahman SM (2015) Using the Mercy Method for weight estimation in Indian children. *Global Pediatric Health*: 1-7. [Crossref]
15. Dicko A, Alhousseini ML, Sidibé B, Traoré M, Abdel-Rahman SM (2014) Evaluation of the Mercy weight estimation method in Ouelessebougou, Mali. *BMC Public Health* 14: 270. [Crossref]
16. Clark MC, Lewis RJ, Fleischman RJ, Ogunniyi AA, Patel DS, et al. (2016) Accuracy of the Broselow Tape in South Sudan, "The Hungriest Place on Earth". *Academic Emergency Medicine*. 23: 21-28. [Crossref]

17. Nieman CT, Manacci CF, Super DM, Mancuso C, Fallon WF Jr (2006) Use of the Broselow tape may result in the underresuscitation of children. *Acad Emerg Med* 13: 1011-1019. [Crossref]

18. Tanner D, Negaard A, Huang R, Evans N, Hennes H (2016) A prospective evaluation of the accuracy of weight estimation using the Broselow Tape in overweight and obese pediatric patients in the Emergency Department. *Pediatric Emergency Care*. [Crossref]

19. Young KD, Korotzer NC (2016) Weight Estimation Methods in Children: A Systematic Review. *Ann Emerg Med* 68: 441-451. [Crossref]

20. Abdel-Rahman SM, Paul IM, James LP, Lewandowski A (2013) Evaluation of the Mercy TAPE: performance against the standard for pediatric weight estimation. *Annals of Emergency Medicine* 62: 332-339.e6. [Crossref]

21. Abdel-Rahman SM, Ahlers N, Holmes A, Wright K, Harris A, et al. (2013) Validation of an improved pediatric weight estimation strategy. *Journal of Pediatric Pharmacology and Therapeutics*. 18: 112-121. [Crossref]

22. Abdel-Rahman A, Jacobsen R, Watts J, Doyle S, O'Malley D, et al. (2015) Comparative performance of pediatric weight estimation techniques: a human factor errors analysis. *Pediatric Emergency Care*.

23. Georgoulas V, Wells M (2016) The PAWPER tape and the Mercy Method outperform other methods of weight estimation in children in South Africa. *South African Medical Journal.* 106: 933-939.

24. Wells M, Coovadia A, Kramer E, Goldstein L (2013) The PAWPER tape: A new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. *Resuscitation* 84: 227-232. [Crossref]

25. Goldstein LN, Wells M (2016) Fat is the new normal. *J Pediatr* 168: 257. [Crossref]

26. Luten R, Zaritsky A (2008) The sophistication of simplicity... optimizing emergency dosing. *Academic Emergency Medicine* 15: 461-465. [Crossref]

27. Kleinman ME, Chameides L, Schexnayder SM, Samson RA, Hazinski MF, et al. (2010) Pediatric advanced life support: 2010 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. *Pediatrics* 126: e1361-99. [Crossref]

28. Wells M, Kramer E (2008) Optimizing emergency drug dosing in children. *Acad Emerg Med* 15: 1325. [Crossref]

29. Srinivasan V, Nadkarni VM, Helfaer MA, Carey SM, Berg RA, American Heart Association National Registry of Cardiopulmonary Resuscitation I (2010) Childhood obesity and survival after in-hospital pediatric cardiopulmonary resuscitation. *Pediatrics* 125: e481-8.

30. Wells M, Goldstein L, Bentley A (2017) High-tech adjuncts to emergency weight estimation: Point-of-care ultrasound and point-of-care bioelectrical impedance measurements can increase the accuracy of length-based weight estimation in children. *Trauma and Emergency Care* 2: 1-8.

Copyright: ©2017 Wells M. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

7.3 Epilogue

7.3.1 Major findings

This was the first study to describe the functioning and performance of the PAWPER XL tape. It was also the first study to formally evaluate the ability of a weight estimation system to estimate IBW as well as TBW.

The PAWPER XL tape estimated TBW more accurately than the Broselow tape and the Mercy method in the study sample. This remained the case in the analysis of the subgroups for which the PAWPER XL tape was designed: children taller than 153 cm and obese children. The PAWPER XL tape was also, therefore, more accurate than the original PAWPER tape.

The PAWPER XL tape estimated IBW very accurately in obese children and was slightly more accurate than the Broselow tape (which itself was very accurate at estimating IBW).

Although this was a preliminary study with relatively small numbers of obese children and children >153 cm, the results were encouraging. The accuracy of the tape system in severely obese children was still lower than anticipated, however. The possibility that there is a limit to the maximum accuracy of this methodology must be considered, either because of the inherent biological variability in the relationship between the discriminant variables, or because of limitations in visual assessment of habitus. Further research will need to establish if weight estimation in this subgroup could potentially be improved with recalibration of the upper habitus score weight values on the tape.

7.3.2 This study in the context of the thesis

The central theme of this thesis is the further development, refinement and then validation of the PAWPER system, in its evolving forms. This study described the first step in the further development of the tape and the initial improvement on the limitations of the PAWPER system that had been identified.

The findings that the PAWPER XL tape was both more accurate than the original PAWPER tape (because it was more accurate in obese children) and that it could provide a weight estimation in all children enrolled in the study (up to the age of 16 years) was testament to the improvement in the system.

The finding that the PAWPER XL tape could predict IBW was useful in the broader context of the fact that weight estimation is needed to enable accurate drug dose determinations. While there is a reasonable consensus on how drugs should be dosed in obese children, the actual processes involved have not been adequately explored. Would it be too complex too estimate both TBW and IBW during an emergency? Would it be too difficult to make the decisions about which weight descriptor to use and too difficult to perform the calculations? Perhaps most importantly, does it matter how drug doses are calculated in obese children? The only available evidence to this last question suggests that it is important and that incorrect drug doses may have a significant effect on outcomes [11, 12]. The answers to the questions of complexity are important and will be covered in Chapter 10.

7.3.3 Link to the next article

This study addressed two of the issues raised by previous studies and provided some preliminary evidence to support the new PAWPER XL tape's design and functioning in taller children and obese children. The subjectivity of habitus assessment and the lack of a habitus-assessment method that could be standardised across different populations had not been addressed by the new tape, however. Since the correct assignment of habitus is essential to preserve the accuracy of the PAWPER system, this needed to be addressed.

The visual assessment of habitus aided by reference images has been described – and validated – in the somatotyping literature as well as the psychological literature for many years [13-15]. It has also been extensively studied in the veterinary field for animal body condition scoring [16]. This provided the inspiration for the development of figural reference images that could be used for training and in clinical practice

together with the PAWPER XL tape. The next chapter describes a study in which the usefulness of reference images in guiding the assessment of children's habitus was compared with the existing standard method. If this system proved accurate, it could increase both the objectivity and standardisation of habitus assessment by different users and in different populations.

7.3.4 References

[1] Young KD, Korotzer NC. Weight estimation methods in children: a systematic review. Annals of Emergency Medicine. 2016;68(4):441-451 e10.

[2] Chavez H, Peterson R, Lo K, Arel M. Weight estimation in an inner city Pediatric Emergency Department: the effect of obesity. American Journal of Emergency Medicine. 2015;33(10):1364-1367.

[3] Garcia CM, Meltzer JA, Chan KN, Cunningham SJ. A validation study of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape-a new weight estimation tool. Journal of Pediatrics. 2015;167(1):173-177.e1.

[4] O'Leary F, John-Denny B, McGarvey K, Hann A, Pegiazoglou I, Peat J. Estimating the weight of ethnically diverse children attending an Australian emergency department: a prospective, blinded, comparison of age-based and length-based tools including Mercy, PAWPER and Broselow. Archives of Disease in Childhood. 2017;102(1):46-52.

[5] Downing A, Rudge G. A study of childhood attendance at emergency departments in the West Midlands region. Emerg Med J. 2006;23(5):391-3.

[6] Baker C. Accident and Emergency statistics: demand, performance and pressure. London: House of Commons Library, 2017.

[7] Twarog JP, Politis MD, Woods EL, Daniel LM, Sonneville KR. Is obesity becoming the new normal? Age, gender and racial/ethnic differences in parental misperception of obesity as being 'About the Right Weight'. International Journal of Obesity (London). 2016;40(7):1051-1055.

[8] Goldstein L, Wells M. Fat is the new normal. Journal of Pediatrics. 2016;168:257.

[9] Wells M, Goldstein L. Optimising emergency weight estimation in underweight and obese children: the accuracy of estimations of total body weight and ideal body weight. European Journal of Emergency Medicine. 2018; In Press.

[10] Rowe S, Siegel D, Benjamin DK, Jr., Best Pharmaceuticals for Children Act - Pediatric Trials Network Administrative Core C. Gaps in drug dosing for obese children: a systematic review of commonly prescribed emergency care medications. Clinical Therapeutics. 2015;37(9):1924-1932.

[11] Srinivasan V, Nadkarni VM, Helfaer MA, Carey SM, Berg RA, American Heart Association National Registry of Cardiopulmonary Resuscitation I. Childhood obesity and survival after in-hospital pediatric cardiopulmonary resuscitation. Pediatrics. 2010;125(3):e481-8.

[12] Pinchevsky LE, Pesaturo KA, Smith BS, Hartman CA. Pilot comparison of three cardiopulmonary resuscitation medication dosing strategies in overweight children. Journal of Pediatric Pharmacology and Therapeutics. 2010;15(4):282-9.

[13] Tanner JM. Reliability of anthroposcopic somatotyping. American Journal of Physical Anthropology. 1954;12(2):257-65.

[14] Gardner RM, Brown D. Body image assessment: A review of figural drawing scales. Personality and Individual Differences. 2010;48:107-111.

[15] Collins M. Body figure perceptions and preferences among preadolescent children. International Journal of Eating Disorders. 1991;10(2):199-208.

[16] Vieira A, Brandao S, Monteiro A, Ajuda I, Stilwell G. Development and validation of a visual body condition scoring system for dairy goats with picture-based training. Journal of Dairy Science. 2015;98(9):6597-608.

Chapter 8

The use of body habitus reference images improves weight estimation using the PAWPER XL tape system.

8.1 Prologue

8.1.1 Background

There is evidence, from two previous studies in international populations, that subjectivity of habitus assessment may have affected the accuracy of weight estimation by the PAWPER tape [1, 2]. Since many users of the tape may not be experienced in assessing habitus using only a visual or gestalt method of assessment, it was important to develop and evaluate other methods to guide training in, and the clinical use of, habitus assessment.

The visual assessment of a child may be affected by the assessor's perception of how a normal weight (non-underweight, non-overweight) child should appear: if the "average" or "standard" child in one population has a higher BMI than in another population, the habitus rating of the "normal" child will differ [3, 4]. This is not the only bias, however. Another problem previously noted with the assessment of habitus – in another context – is the visual bias associated with recognising obesity [5]. The main element of this bias that is relevant to visual habitus assessment is contraction bias. This bias refers to the natural tendency for an observer to underestimate the size of visual stimuli (objects) that are larger than average [6]. With respect to overweight and obese children, this means that their degree of obesity (or habitus score) will be prone to underestimation and this underestimation will be magnified as their BMI increases [5-7]. Unaided visual assessment is, therefore, unlikely to be completely reliable and other methods to improve its accuracy must be found.

Visual assessment of habitus has been studied in other fields. Body image studies have used and validated figural reference images against BMI [8-10]. Nutritional studies and body composition studies have used silhouettes to identify underweight and obese children [11-13]. Veterinary studies have used similar images to quantify the nutritional status of domestic and wild animals [14-16]. In the light of these

precedents, it was highly likely that a similar system could be effective to aid the appraisal of habitus for use with the PAWPER XL tape.

Checklists have been shown to play an important role in many forms of decision-making. For this reason, it was essential to evaluate the potential role of checklists in the assessment of habitus. A specifically designed checklist was evaluated as another method of assessing habitus in this study.

8.1.2 Purpose of the study

A preliminary study on using the modified Collins figures had proven successful at improving the habitus classification of obese children when using the PAWPER XL tape [17]. Two international studies, which only used images of very young children for training in habitus assessment were not as successful [1, 2]. This is probably because the ages of the children portrayed in the references images must be similar to the age of the child being assessed [18]. For this reason, a set of images was created for infants, young children (aged 4 to 5 years) and older children (age 10 to 12 years) to allow the user to select the appropriate reference image for use.

The main purpose of this study was to investigate the possibility of using adjuncts to assist anthroposcopic habitus classification. The study compared gestalt-based, checklist-assisted and reference image-assisted methods of habitus assessment to establish which method was the most objective and most accurate.

Since these methods were intended to be used during emergency medical care, we also wanted to investigate which method was the most satisfactory to the users. The participants' opinions were canvassed as to which method they thought would be most useful and easiest to use during an emergency. Since the method recommended for future use with PAWPER XL tape would be dependent on its acceptance by the users, this was an important objective of this study.

This study provided data for Objective 4 (to establish whether the accuracy of weight estimation by the PAWPER XL method could be improved by the use of figural reference images to aid in the assessment of body habitus).

8.2 Published paper

Wells M, Goldstein L, Bentley A. The use of body habitus reference images improves the ability of novices to accurately estimate children's weight using the PAWPER XL tape system. The Journal of Emergency Medicine. 2017;54(2):165-175.

The PDF of the published article is included. The online-only supplementary material is appended after the PDF. The tables or figures have been retained in the formatting in which they are archived in the journal website.

The Journal of Emergency Medicine, Vol. 54, No. 2, pp. 165–175, 2018
© 2017 Elsevier Inc. All rights reserved.
0736-4679/$ - see front matter

https://doi.org/10.1016/j.jemermed.2017.10.009

Original Contributions

THE USE OF BODY HABITUS REFERENCE IMAGES IMPROVES THE ABILITY OF NOVICES TO ACCURATELY ESTIMATE CHILDREN'S WEIGHT USING THE PAWPER XL TAPE SYSTEM

Mike Wells, MBBCH, MSC, FCEM(SA), Lara Nicole Goldstein, MBBCH, MMED, FCEM(SA), and Alison Bentley, MBBCH, PHD

Division of Emergency Medicine, Faculty of Health Sciences, University of the Witwatersrand, Parktown, Johannesburg, South Africa
Reprint Address: Mike Wells, MBBCH, MSC, FCEM(SA), Postnet Suite 429, Private Bag X1510, Glenvista 2058, South Africa

□ **Abstract—Background: The original methodology of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape relies on a gestalt visual assessment of a child's body habitus to adjust a length-based weight estimation. This assessment is dependent on the user's subjective opinion, which may result in aberrations in accuracy between users and populations. With the development of the second-generation PAWPER XL tape, a more objective method of habitus evaluation was desired. Objective: The aim of this study was to evaluate a new, more objective figural reference image system and a new checklist system for quantifying body habitus. Methods: Volunteers were asked to assess the body habitus score of 90 children from a sequence of photographic images using the standard gestalt visual assessment system, a new checklist system, and a system using figural reference images. PAWPER XL tape weight estimations were generated from these scores, which were compared between the three test groups. Participants were also surveyed on their preferences for the different methodologies. Results: There were 11,505 habitus score assessments from 138 doctor, nurse, and paramedic participants. The figural reference image system significantly outperformed the checklist system and the gestalt visual assessment system in terms of weight estimation accuracy, achieving 70.9%, 61.1%, and 60.9% of estimations within 10% of measured weight, respectively. The participants expressed a strong subjective preference for the image system because of speed of use, ease of use, perceived accuracy, objectivity, and low cognitive load. Conclusions: The figural reference image system was objectively much more accurate than, and subjectively preferable to, the original gestalt visual estimation methodology.** **© 2017 Elsevier Inc. All rights reserved.**

□ **Keywords—body weight; resuscitation; weight estimation; PAWPER tape**

Conflict of Interest: Mike Wells is the developer of the PAWPER XL tape.

INTRODUCTION

When weight cannot be measured during the management of medical emergencies in children, an accurate estimation of total body weight is required to facilitate the delivery of safe and effective doses of medications (1). Significant errors in weight estimation may result in large medication dosing errors, with the concomitant potential for harm, especially in obese and underweight children (2,3). It is therefore imperative to use the most accurate and appropriate method available for emergency weight estimation: the use of a system that is known to be inaccurate cannot be considered to be good medical practice (4,5).

The PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape is one of the dual length- and habitus-based methods specifically designed to be used during emergency care (the use of the PAWPER tape and the habitus scores are described in Figure 1)

RECEIVED: 17 June 2017; FINAL SUBMISSION RECEIVED: 5 September 2017;
ACCEPTED: 7 October 2017

(6,7). It has been found to be extremely accurate in most populations, but somewhat less accurate in populations with a very high prevalence of obesity (6–11). One of the main reasons for this reduced accuracy was user-error—an underestimation of fatness in overweight and obese children leading to a suboptimal assessment of habitus (10). The PAWPER tape methodology relies on the correct identification of a child close to the 50th percentile of weight for length as the "average" child. Because the perception of the "average" child may differ between populations, especially in populations where "fat is the new normal," this could lead to errors (12).

Although there is evidence that visual assessment of body habitus for the purposes of weight estimation can be accurate and repeatable, these assessments could also be less reliable in raters of different skill or experience (13,14). The use of a more objective, standardized method to assess body habitus would, therefore, be of potentially great benefit, as it could increase the accuracy of the resultant weight estimation by the next generation of the PAWPER tape, the PAWPER XL tape.

The aim of this study was to evaluate the use of objective, non-anthropometric methods of assessing body habitus to establish whether they could improve the accuracy of weight estimates produced by the PAWPER XL tape. The specific objectives of this study were to investigate whether the use of a checklist-based system or a figural reference image-based system to assess habitus would improve the resulting accuracy of weight estimation when compared with the current standard system (gestalt-based visual assessment).

METHODS

Development of Checklist and Figural Reference Image Systems

We created a checklist to evaluate and quantify body habitus based on easily and rapidly observable anatomic features (Figure 2). It was loosely modeled on existing animal condition scoring systems, anthroposcopic somatotyping techniques, and descriptions used in the development of figural rating scales (15–17). We developed the body habitus reference image model for this study based on existing images, including the Stunkard figural rating scale, the Collins scale, and the body image assessment scale (18–20). We used the validated body mass index (BMI) scores of these systems to "calibrate" and stratify the images into habitus score categories. A professional artist adapted and created images for three different age groups in frontal and profile views (Figure 2 and Supplementary Figure 1).

Ethics

We obtained ethics approval for this study from the Human Research Ethics Committee of the University of the Witwatersrand. Each participant consented to participation in the study and we acquired consent to use the

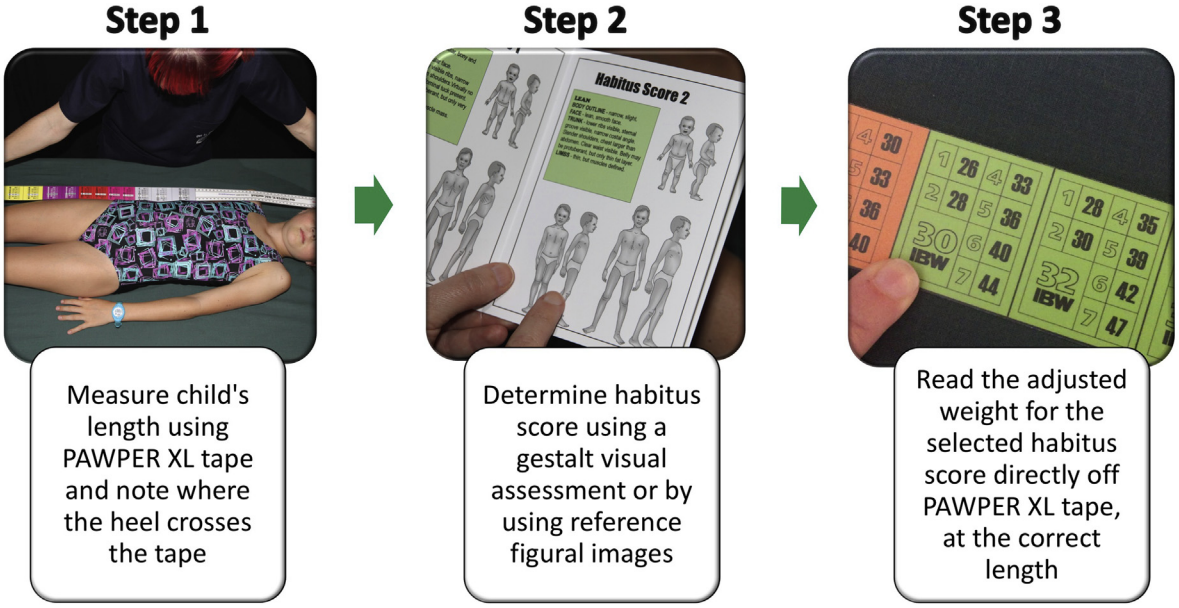


Figure 1. How the updated version of the original PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape, the PAWPER XL tape, is used. While the measurement of length is objective, the assessment of habitus score may not be. The original tape methodology called for the user to assign a habitus score based on a simple visual estimation of habitus, which can be very subjective. The figure shows the enhanced body habitus reference image system that was devised after this study.

HABITUS SCORING FOR THE PAWPER XL TAPE	
1 THIN	**Body outline -** Angular, bony and narrow. **Face -** Thin, angular face. **Chest -** Clearly visible ribs. Thin shoulders. **Abdomen -** Virtually no fat. Flat, abdominal tuck present. Belly may be protuberant, but only very thin fat layer. **Limbs –** Thin. Very poor muscle mass.
2 LEAN	**Body outline -** Narrow, slight. **Face -** Lean, smooth face. **Chest -** Lower ribs visible, sternal groove visible. Slender shoulders. Chest larger than abdomen. **Abdomen -** Flat, clear waist present. Belly may be protuberant, but only thin fat layer. **Limbs -** Thin but muscles defined.
3 OPTIMUM	**Body outline -** Smooth outline. **Face -** Smooth, rounded face. **Chest -** Outline of ribs just visible; no sternal groove. Chest equal to abdomen in size. **Abdomen -** Smooth layer of fat. Flat, small waist. **Limbs –** Well-proportioned limbs. Good muscle mass.
4 SLIGHTLY OVERWEIGHT	**Body outline –** Rounded. **Face -** Broad, chubby face. **Chest -** Smooth fat layer over ribs, no ribs visible, or just visible. Broad shoulders. Abdomen larger than chest. **Abdomen -** Thick layer of fat. Rounded abdomen, high waist, no waist or very minimal waist. **Limbs -** Noticeable thickening of limbs.
5 OVERWEIGHT	**Body outline -** Bulges from fat. **Face -** Rounded, fleshy face, big cheeks, some fat on neck. **Chest -** Fatty "breasts" visible. Broad chest. **Abdomen -** Thick layer of fat, with rolls and folds. High waist / no waist, rounded abdomen. Abdomen much larger than chest. **Limbs -** Thick fat layer over limbs.
6 OBESE	**Body outline -** Very rounded, lumps of fat. **Face -** Rounded face, small eyes, rolls of fat on neck. **Chest -** Thick layer of fat with distinct fatty "breasts". Very broad shoulders. Chest and abdomen rounded together. **Abdomen -** Rolls and lumps of fat. No waist, protuberant abdomen. **Limbs -** Fat rolls and lumps on proximal limbs.
7 SEVERELY OBESE	**Body outline –** Rotund with bulges of fat. **Face -** Rounded, heavy jowls, thick fat rolls on neck. **Chest -** Large, fatty "breasts" with folds and rolls. Both chest and abdomen protuberant from thick fat layer. **Abdomen -** Pendulous fat layer. No waist, massively rounded. **Limbs -** Thick layer of fat covering limbs; pronounced "hamming".

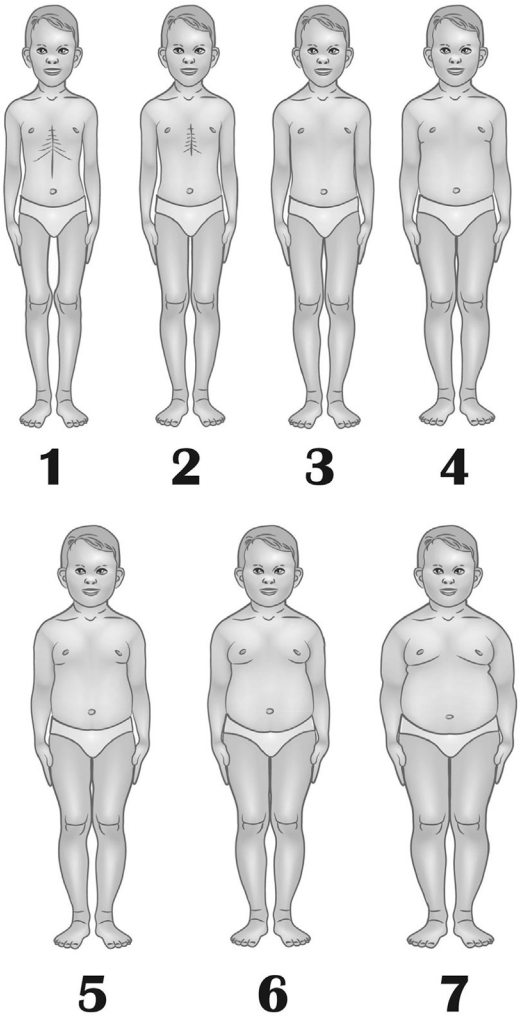


Figure 2. Checklist criteria and examples of the figural reference images used in this study. The criteria included in the checklist are shown in the left panel and an example of the frontal views for children aged 5 to 10 years. Three sets of figural stimuli were used: for ages < 5 years, ages 5 to 10 years and age > 10 years (see Supplementary Figure 1). The checklists and images were on different pages and not available to the raters at the same time. PAWPER = Pediatric Advanced Weight Prediction in the Emergency Room.

anonymized photographs of the children rated by the participants.

Study Design and Sample

We conducted a prospective, crossover trial with volunteers from four groups of health care providers: doctors in full-time emergency medicine practice, pediatricians or pediatric residents, professional nurses, and advanced life support paramedics. We ran six educational sessions between February and August 2015, during which time we recruited the volunteers. Before beginning the study, the participants completed a questionnaire to provide basic demographic information. We excluded volunteers

who had used the PAWPER tape more than once previously in clinical practice from the study, to reduce potential bias introduced by experience.

Sample Size

We estimated that a minimum sample size of 120 participants (25 in four discipline subgroups) was required to detect a 10% difference in accuracy between the habitus assessment methods (Fisher exact test), powered to 80% at a 0.05 significance level.

Study Procedure

We showed the participants a slideshow sequence of frontal photographs of children and infants aged from 6 months to 16 years, representative of the full spectrum of body habitus types (very underweight to morbidly obese), obtained from a previous study (8). The children were clothed in underwear and had their faces covered for anonymity. We asked the participants to examine the images and record their assessment of the habitus score (HS) from HS1 to HS7 on a pro forma data-collection sheet, with HS1 representing an underweight child, HS3 the "average" child, HS5 an obese child, and HS7 a morbidly obese child (Figure 3).

The participants assessed 90 images of children in three sequences, with a short break between each: they rated 30 images using the gestalt visual assessment method; 30 using a checklist method; and 30 using a figural reference image method. Before each sequence, we conducted a 15-min training session on the rating method that the participants were to use next. We selected the images so that each test group's image sequence was matched for age, weight, and body habitus. We displayed each image for only 15 s, to simulate real-world use and to simulate minor cognitive stress. We devised the sequence in which the assessment methods were evaluated to minimize the potential bias induced by learning: visual gestalt assessment was evaluated first, then the checklist method, and finally, the image-assisted method.

At the end of the study, we asked the participants to complete another questionnaire to obtain open-ended feedback on which of the systems they preferred, which they found easiest to use, and which they perceived as the most accurate.

Data Analysis

We calculated the estimated weight for each child using the PAWPER XL tape method from the habitus score ratings recorded by the participants and the known length of each child used in the photographs. We evaluated the accuracy of these weight estimations by calculating the percentages of estimations falling within 10% and 20% of the actual weight of the child (PW10 and PW20). We

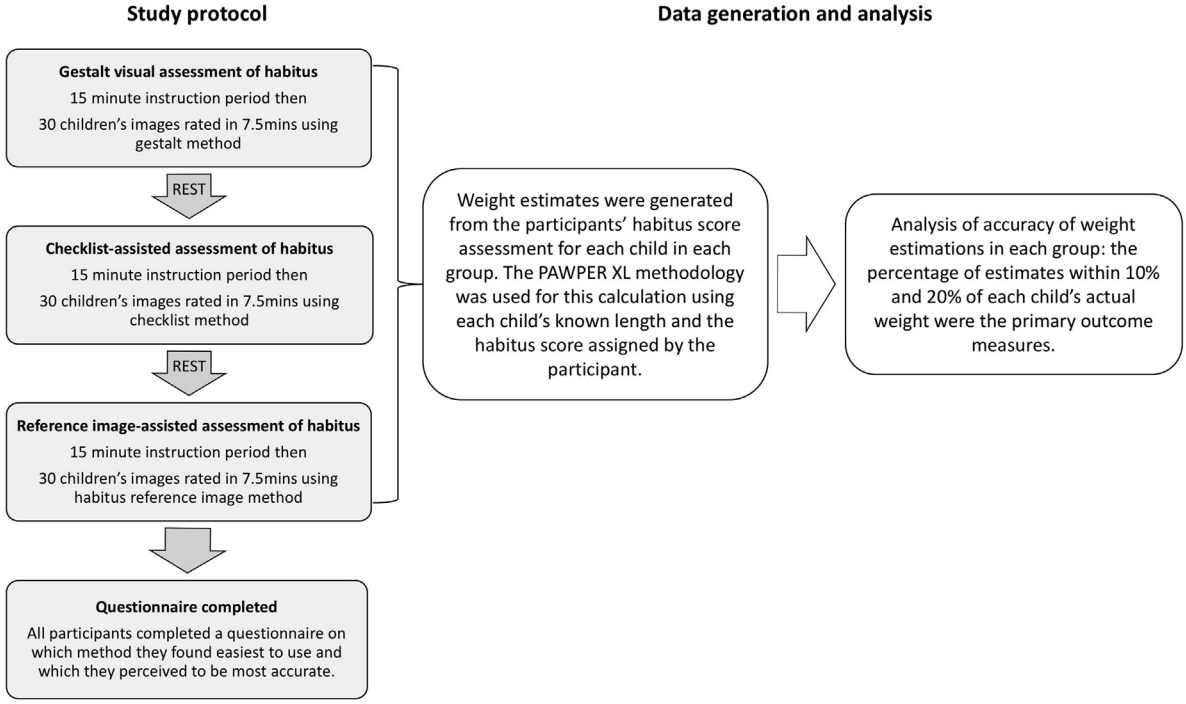


Figure 3. Flow chart showing the study procedure. PAWPER = Pediatric Advanced Weight Prediction in the Emergency Room.

defined an acceptable level of accuracy as a PW10 ≥ 70% and a PW20 ≥ 95% (4). We calculated the accuracy of weight estimates for the best and worst 20 raters, as well as for the best and worst 50% of the participants within each test group. To provide a means of evaluating the performance of the three systems against other weight estimation methods, we also presented the accuracy of weight estimation obtained using the Broselow tape and the Mercy method for the study sample.

We calculated the correctness of HS assessment for each rater and for each of the three test groups. We also evaluated the inter-rater reliability within each group using intra-class correlation.

We made comparisons of accuracy data between groups using the Fisher exact test. In addition, we performed subgroup analyses by age (< 5 years, 5–10 years, and > 10 years) and habitus (underweight, "normal" weight, and obese).

We performed a multiple regression analysis to identify any associations between demographic characteristics and weight estimation accuracy outcomes.

Finally, we analyzed the qualitative data to identify themes from the participants' feedback on the strengths and weaknesses of each of the habitus-assessment systems tested.

Outcome Measures

The primary outcome measure was if a habitus scoring system achieved an acceptable level of weight estimation accuracy (i.e., a PW10 > 70% and a PW20 > 95%). The secondary outcome measures were an improvement in overall weight estimation accuracy of the new systems of ≥ 10% over the standard visual estimation technique, an improvement in the accuracy of HS assessment of ≥ 10% over the standard visual assessment technique, and a higher level of user satisfaction with the new systems when compared to the standard gestalt method.

Software

We performed the statistical analysis with StatSoft, Inc. (2011, STATISTICA data analysis software system, version 10, www.statsoft.com).

RESULTS

Characteristics of Participants and Children Models

Of the 141 participants enrolled, 3 were excluded because of previous experience with the use of the PAWPER tape, resulting in a total number of 138 raters who produced 11,505 habitus assessments. The demographic characteristics of the participants and the children comprising each group are shown in Table 1. There were no significant differences between the characteristics of the children in the three assessment groups.

Accuracy of Weight Estimations

The accuracy of the PAWPER XL weight estimations derived from the HS of each system of habitus assessment are shown in Figure 4. The reference image method achieved the acceptable outcome benchmark in the overall analysis and in every subgroup except when used in obese children and when used by nurses. The gestalt method failed to achieve the benchmark in any subgroup, while the checklist method achieved it only in the subgroup of children with average habitus (Table 2). Of the 128 participants, 119 (86.2%) raters achieved their highest accuracy with the reference image method, 12 (9.4%) were most accurate with the checklist system and only 7 (4.3%) with the gestalt method. Nurses performed significantly worse with the image system than doctors and paramedics (Fisher exact test, $p = 0.0025$ and $p = 0.0046$, respectively).

The multiple regression analysis did not show any significant associations between weight estimation accuracy and health care provider type, experience, number of children seen per week, or confidence levels.

There was substantial variation in the performance between individual raters (see the best and worst performance data in Table 2). Both the best and worst individual raters had better outcomes with the reference image method.

The accuracy of weight estimation by the Broselow tape and the Mercy method in the sample of children rated by the participants is shown in Table 2.

Accuracy of HS Assessment

The outcomes of the analysis of correctness of habitus assessment are shown in Supplementary Figure 2. The reference image method produced a significantly higher number of correct HSs than the other two methods in the overall sample, as well as many of the subgroups. The gestalt system produced a substantially larger number of underestimations of HS than the checklist and reference image systems (47.3%, 36.5%, and 29.0%, respectively, Fisher exact test, $p < 0.001$ for each comparison). With regard to the best and worst raters, a very similar pattern to the weight estimation data emerged. The inter-rater reliability was similar between the groups, and could be classified as "good" (intra-class correlation 0.6 to 0.74).

Table 1. Demographic Characteristics of the Participants and Children Making Up the Three Test Groups

Variable	n	Sex, Male, n/N (%)	Experience (Years), Median (UQ, LQ) (n)	No. of Children Seen per Week, Median (UQ, LQ)	Confidence, Median (UQ, LQ)
Details of participants (raters)					
All	138	59/134 (44.0)	6.0 (3.0, 9.8) (102)	4.0 (1, 11)	5.0 (4.0, 6.0)
All doctors	80	36/80 (46.2)	6.0 (3.0, 9.0) (63)	1.0 (1, 10)	5.0 (4.0, 6.0)
Pediatrics	27	7/27 (26.9)	6.0 (2.5, 9.0) (26)	35 (18, 50)	4.5 (2.8, 7.0)
Emergency medicine	53	23/43 (55.8)	6.0 (3.0, 9.0) (37)	1 (0.5, 4)	5.0 (4.0, 6.0)
Nurses	25	5/25 (25.0)	9.0 (7.5, 20) (23)	10 (5, 18)	4.0 (4.0, 5.0)
ALS paramedics	32	18/32 (56.3)	4.0 (4.0, 7.0) (16)	2 (1, 2)	4.5 (3.0, 6.0)

Variable	All (n = 90)	Gestalt Group (n = 30)	Checklist Group (n = 30)	Image Group (n = 30)	*p* Value
Details of children					
Age, median (LQ, UQ)	7.3 (5.4, 9.7)	8.0 (5.6, 9.7)	6.6 (5.3, 9.6)	7.9 (5.8, 10.4)	NS
Weight, median (LQ, UQ)	25.2 (19.8, 36.4)	25.7 (20.5, 38.2)	21.4 (18.6, 34.2)	25.3 (20.0, 33.9)	NS
IBW, median (LQ, UQ)	24.0 (19.0, 32.0)	26.0 (20.0, 33.5)	24.0 (19.0, 27.5)	26.0 (19.8, 32.0)	NS
Length, median (LQ, UQ)	1.24 (1.11, 1.37)	1.28 (1.15, 1.39)	1.21 (1.1, 1.27)	1.25 (1.11, 1.37)	NS
BMI, median (LQ, UQ)	17.2 (15.1, 19.6)	17.3 (15.4, 19.4)	16.6 (14.6, 19.7)	17.2 (15.7, 19.6)	NS
BMI-for-age *Z* score, median (LQ, UQ)	0.6 (−0.3, 1.8)	0.5 (−0.3, 2.0)	0.7 (−0.8, 1.8)	0.6 (−0.3, 1.5)	NS
Body fat, %, median (LQ, UQ)	31.0 (25.0, 37.0)	34.5 (24.0, 39.0)	31.0 (26.0, 37.0)	30.0 (26.3, 36.0)	NS
Underweight, *Z* score < −1.4, n (%)	14 (15.6)	4 (13.3)	6 (20.0)	4 (13.3)	NS
Overweight, 1.4 < *Z* score < 2.0, n (%)	24 (26.7)	8 (26.6)	8 (26.6)	8 (26.6)	NS
Obese, 1.9 < *Z* score < 2.5, n (%)	11 (12.2)	3 (10.0)	5 (16.7)	3 (10.0)	NS
Morbidly obese, *Z* score > 2.4, n (%)	8 (8.9)	3 (10.0)	2 (6.7)	3 (10.0)	NS
Habitus score, median (LQ, UQ)	3.0 (2.0, 5.0)	3.0 (3.0, 4.0)	3.0 (2.0, 5.0)	3.0 (2.0, 5.0)	NS

ALS = advanced life support; BMI = body mass index; IBW = ideal body weight; LQ = lower quartile; UQ = upper quartile.
There were no significant differences in anthropometric and body composition data between the children used to make up each test group. Underweight, overweight, obese, and morbidly obese children were defined according to BMI-for-age *Z* scores of < −1.0, >1.4, >2, and >2.5, respectively. Body fat percentage was available from dual-energy x-ray absorptiometry measurements.

Participant Preferences

The participants' responses to their experiences with the three systems are shown in Table 3. Two-thirds of the participants preferred the image method to the other systems. The reference image method was strongly supported with positive comments about accuracy, ease of use, objectivity, and speed of use. There were very few negative comments. The participants felt the gestalt method was easy to use, but they were worried about guessing incorrectly when assessing habitus.

DISCUSSION

Weight Estimation Accuracy

The weight estimation accuracy achieved in this study using the figural reference image system was very good, considering this was accomplished by novices in a sample in which > 60% of children were underweight, overweight, or obese. To provide perspective, the accuracy achieved was substantially higher than that produced by the Broselow tape or the Mercy method (one of the most accurate weight estimation systems in the world) in this same study sample when used by experts. The figural reference image system was significantly superior to the other methods and it surpassed the primary and secondary outcome measures with respect to acceptable performance in the overall sample and all subgroups except in obese children and when used by nurses.

The image system performed worse in the subgroup of obese children than in average and underweight children, although it was still significantly better than the gestalt test group (odds ratio 5.6). Studies in other fields have also frequently shown that "fatness" is underestimated in children, for reasons that are not clear. It is possible that the increasing prevalence of childhood obesity has meant that "fat is the new normal" and reduced our ability to identify overweight children as such (12,21,22). A "calibration" error of the images must also be considered, although the images used in this study were developed to match other figural reference images that have been validated and calibrated according to their ability to predict BMI (18–20). This needs further research.

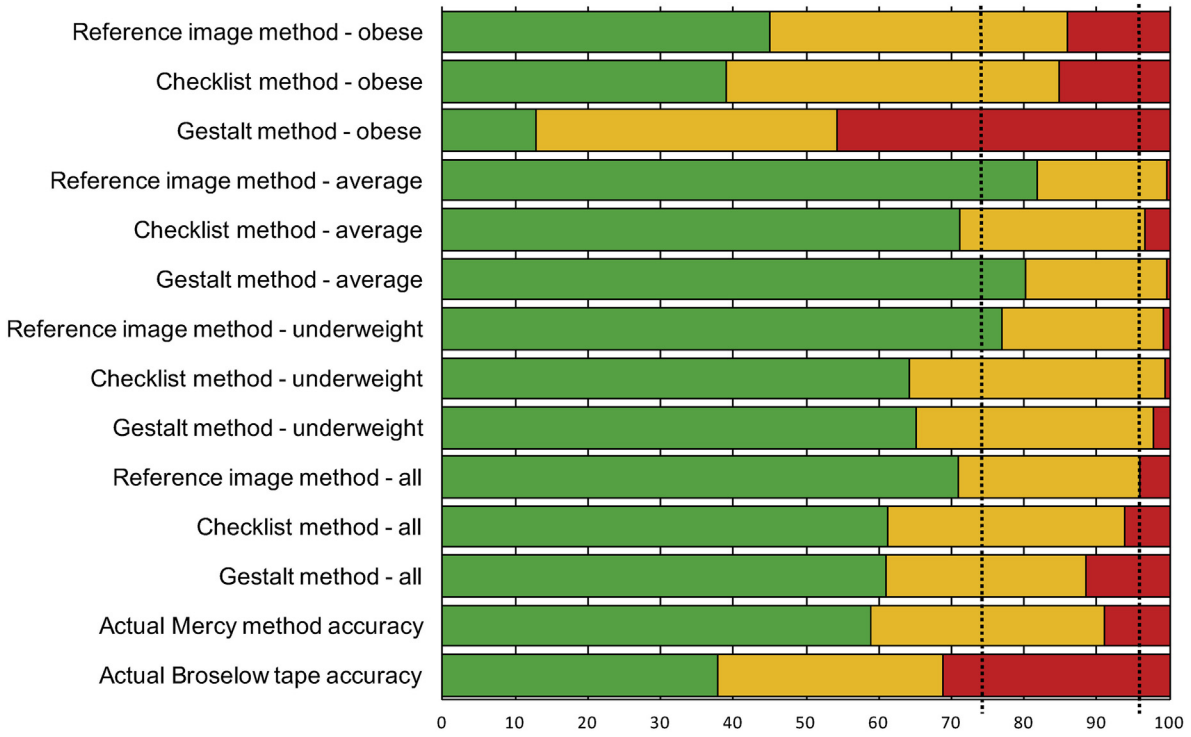


Figure 4. The accuracy of weight estimates based on the original gestalt methodology and the two new methodologies of habitus assessment. The green bars indicate the percentage of weight estimations within 10% of actual weight. The orange plus the green bars indicates the percentage of weight estimations within 20% of actual weight. The red bars indicate critical weight estimation errors (> 20% error). The dotted lines indicate the desirable accuracy outcomes (PW10 > 70% and PW20 > 95%). (A color figure can be found in the online version of this article.)

HS Assessment Accuracy

The superior weight estimation accuracy of the reference image system was, unsurprisingly, from a significantly higher accuracy of HS selection. The reference image system also reduced the number of outlier HSs and thereby overall accuracy. There is limited previous work on any form of habitus assessment for the purposes of weight estimation. There is some evidence that visual assessment is accurate and reliable, although this has not been evaluated stringently in children of different habitus categories (13). The use of figural images for weight estimation has also been shown to be useful in two previous studies. One study on the PAWPER tape in South Africa, using figural reference images derived from Collins figures, showed an excellent performance with an improved accuracy in obese children when compared to an experienced gestalt visual assessment (6). Another Australian study successfully used images to identify children with above-average adiposity (14).

There is precedence in the use of figural reference images in other fields: psychologists have used them to evaluate body image in adults and children for decades and their correlation with BMI has been well validated (23–27). The use of images to make reference comparisons has been shown to be mentally straightforward and reliable, so the use of such a system in emergencies has sound theoretical support (28).

Novices and Training

The overall performance by novices in this study was good, but the best novice raters demonstrated an exceptionally high level of accuracy in weight estimation. This proved the potential accuracy achievable by the system. The worst-performing raters with the reference image system were significantly better than with the other systems, which suggests a higher resiliency (lower vulnerability to errors). Previous work from other fields has shown that experience will improve the accuracy of body habitus assessment by novices by about 25%, and that habitus assessment can become very good with training (17,22). Human-factor error studies on the Broselow tape and the Mercy method have shown a high level of errors even when the simplest systems are used: health care providers are frequently undertrained in the use of weight estimation systems (29,30). However, even very limited training has been shown to improve the use of weight estimation systems considerably (31).

Table 2. Accuracy of the Weight Estimation Data Derived from the Habitus Scores Generated for Each Test Group

		Accuracy, Bias, and Precision of Weight Estimation								
		Gestalt Method		Checklist Method		Reference Image Method		Checklist vs. Gestalt, Relative Change, %; OR (95% CI); *p* Value	Reference Image vs. Gestalt, Relative Change, %, OR (95% CI); *p* Value	Reference Image vs. Checklist, Relative change, %, OR (95% CI); *p* Value
Variable	n	PW10, %	PW20, %	PW10, %	PW20, %	PW10, %	PW20, %			
All	4123	***60.9***	88.5[1]	***61.1***	93.9	70.9	95.9[1]	0.3	16.4; 1.6 (1.4–1.7); <0.0001	16.0; 1.6 (1.4–1.7); <0.0001
Thin (HS < 3)	961	***65.1***	97.7	***64.3***	99.4	76.9	99.1	*1.2*	18.1; 1.8[4] (1.5–2.2); <0.0001	19.6; 1.8[5] (1.5–2.3); <0.0001
Average (HS 3 and 4)	2199	80.1	99.5	71.1	96.5	81.8	99.6	*12.7*; 1.3 (1.1–1.5); 0.0011	2.0	15.0; 1.4 (1.2–1.7); <0.0001
Obese (HS > 4)	963	***12.9***	54.3[2,3]	***38.9***	84.9[2]	***45.1***	85.9[3]	201.6; 4.3 (3.4–5.4); <0.0001	249.6; 5.6 (4.4–6.9); <0.0001	15.9; 1.3 (1.1–1.5); 0.006
Age < 5 years	963	***60.3***	98.9	***65.2***	97.3	70.2	98.4	8.1	12.8; 1.4 (1.2–1.7); 0.0005	4.3
Age 5–10 years	2199	***62.4***	86.6[4,5]	***59.7***	93.5[4]	80.5	96.1[5]	*4.5*	29.0; 2.5 (2.1–2.9); <0.0001	34.8; 2.8 (2.4–3.2); <0.0001
Age >10 years	961	***58.0***	82.5[6,7]	***57.7***	88.2[6,8]	70.3	98.5[7,8]	*0.5*	2.0	4.5
Doctors	2399	***61.3***	89.1[9,10]	***62.5***	94.6[9]	72.5	96.2[10]	2.0	18.3; 1.7 (1.5–1.9); < 0.0001	16.0; 1.6 (1.4–1.8); <0.0001
Nurses	772	***59.2***	86.7[11,12]	***61.6***	93.2[11]	***64.8***	94.5[12]	4.1	9.5; 1.3 (1.0–1.6); 0.02	5.2
ALS	952	***61.3***	88.6[13,14]	***58.3***	93.5[13]	71.7	96.1[14]	*5.1*	17.0; 1.6 (1.3–1.9); <0.0001	23.0; 1.8 (1.5–2.2); <0.0001
Pediatricians	809	***59.7***	89.1[15]	***62.3***	94.2	71.4	96.5[15]	4.4	19.6; 1.7 (1.4–2.1); <0.0001	14.6; 1.6 (1.4–1.9); <0.0001

		Accuracy of Individual Raters								
		Gestalt Method		Checklist Method		Reference Image Method		Checklist vs. Gestalt, Relative Change, %; OR (95% CI); *p* Value	Reference Image vs. Gestalt, Relative Change, %, OR (95% CI); *p* Value	Reference Image vs. Checklist, Relative change, %, OR (95% CI); *p* Value
Variable	n	PW10, %	PW20, %	PW10, %	PW20, %	PW10, %	PW20, %			
Best 20 raters	20	70.3	—	72.7	—	81.2	—	3.4	15.4; 1.8 (1.4–2.4); <0.0001	11.6; 1.7 (1.3–2.2); 0.0002
Worst 20 raters	20	50.2	—	50.6	—	58.2	—	0.8	15.9; 1.4 (1.1–1.7); 0.0064	14.9; 1.4 (1.1–1.7); 0.0107
Best 50% of raters	64	66.0	—	67.7	—	76.7	—	2.6	16.1; 1.7 (1.5–1.9); <0.0001	13.2; 1.6 (1.4–1.8); <0.0001
Worst 50% of raters	64	55.8	—	56.3	—	64.9	—	0.9	16.2; 1.5 (1.3–1.7); <0.0001	15.2; 1.4 (1.3–1.6); <0.0001

Variable	Accuracy of the Broselow Tape and the Mercy Method in This Study Population, as Reference Data		
	n	PW10, %	PW20, %
Broselow tape	90	37.8	68.9
Mercy method	90	58.9	91.1

ALS = advanced life support; CI = confidence interval; HS = habitus score; OR = odds ratio.

PW10 and PW20 refer to the percentage of weight estimations failing within 10% and 20% of actual weight, respectively; these are measures of accuracy. The numbers indicated in bold and italic type indicate values that fell below the predetermined acceptability threshold of accuracy. The PW10 accuracy data are also shown for the best and worst performing participants in each test group. The ORs were calculated from a Fisher exact test and the resulting significance levels are shown. The results of the Fisher exact test between the PW20 outcomes are shown in the bottom panel. The significant outcomes of the statistical comparisons between the PW20 values for each group or subgroup are shown here (Fisher exact test). The superscripts indicate which groups or subgroups were compared and the respective OR (95% CI) and significance level outcomes are shown below.

[1] OR 3.0 (95% CI 2.5–3.6); $p < 0.0001$.
[2] OR 4.7 (95% CI 3.8–5.9); $p < 0.0001$.
[3] OR 5.1 (95% CI 4.1–6.3); $p < 0.0001$.
[4] OR 2.2 (95% CI 1.8–2.8); $p < 0.0001$.
[5] OR 1.9 (95% CI 1.5–2.3); $p < 0.0001$.
[6] OR 1.6 (95% CI 1.1–2.2); $p = 0.0057$.
[7] OR 14.0 (95% CI 8.6–22); $p < 0.0001$.
[8] OR 8.5 (95% CI 5.1–14); $p < 0.0001$.
[9] OR 2.1 (95% CI 1.7–2.8); $p < 0.0001$.
[10] OR 3.1 (95% CI 2.4–3.9); $p < 0.0001$.
[11] OR 2.1 (95% CI 1.5–3.0); $p < 0.0001$.
[12] OR 2.6 (95% CI 1.8–3.8); $p < 0.0001$.
[13] OR 1.8 (95% CI 1.3–2.5); $p = 0.0002$.
[14] OR 3.2 (95% CI 2.1–4.6); $p < 0.0001$.
[15] OR 3.4 (95% CI 2.2–5.3); $p < 0.0001$.

Table 3. Participants' Feedback on the Habitus Scoring Systems*

Variable	Questionnaire Data		
	Gestalt Method	Checklist Method	Image Method
Overall preference by participants, n (%)†	27 (19.6)	14 (10.1)	93 (67.4)
No. of responses on common themes			
Method was quick to use	6	0	31
Method decreased thinking	0	1	14
Method felt highly accurate	2	10	27
Method was objective	0	1	19
Method allows for standardisation	0	1	5
Methodology was distracting	2	8	2
Methodology was time-consuming	0	6	0
I was worried about guessing	13	0	0
I found the method easy to use	19	6	50

* The commonly mentioned themes as well as the indicated overall preference are incorporated.
† No preference: n = 4 (2.9%).

Preferences

Most participants (67.4%) favored the image system, in addition to a significant majority being most accurate with it (86.2%). The image system was regarded as being the easiest to use, providing the most standardization and objectivity, as well as being the most accurate. These perceptions are important because a user's concerns about a system could be an impediment to its use in a real-world scenario. Although the checklist system was also considered to be accurate, participants expressed concerns over it being distracting and time-consuming. Similarly, although the gestalt system was regarded as being easy to use, the respondents were worried about the inaccuracy and stress of having to guess. On the contrary, the checklist system would require slow, analytical thinking, which would be inappropriate during resuscitative care.

Limitations

There were several important limitations to this study. Photographs rather than real-life children were used for assessing HS, which may have been more difficult. This will require validation of the findings in a more clinical setting in the future.

The sequence of assessments may have favored the reference image system, as it was the last method evaluated. Learning and accommodation to stress may have occurred through the study, with increased downstream accuracy of habitus assessment. The study was designed so that each downstream sequence introduced different, nontransferable skills to minimize the effects of learning. The testing of gestalt visual assessment after the checklist or image-assisted method would have induced a greater potential bias due to learning. However, the uncertainty of the possible effects of learning remains a limitation.

The calculation of weights from each child's known length and the HSs generated in the study, rather than from the actual use of the tape on a child, may have introduced bias. This was used in the same way for each of the groups, however, and is unlikely to have made a significant difference in the results.

CONCLUSIONS

The body habitus reference image system was objectively much more accurate than both the original gestalt visual estimation method and the checklist method. It was also the system most subjectively favored by the participants. We concluded that this system would be the optimum method of habitus assessment to use with the PAWPER XL tape.

Acknowledgments—Liesl Roux from Fyrebird creations created the figural reference images.

Author contributions: Mike Wells conceptualized and designed the study, carried out the initial data collection, drafted the initial manuscript, and approved the final manuscript as submitted. Lara Nicole Goldstein assisted with initial data collection, drafted revisions of the manuscript, and approved the final manuscript as submitted. Alison Bentley assisted with preparation of the manuscript and approved the final manuscript as submitted. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

SUPPLEMENTARY DATA

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jemermed.2017.10.009.

REFERENCES

1. Luten R, Wears RL, Broselow J, Croskerry P, Joseph MM, Frush K. Managing the unique size-related issues of pediatric resuscitation: reducing cognitive load with resuscitation aids. Acad Emerg Med 2002;9:840–7.
2. Pinchevsky LE, Pesaturo KA, Smith BS, Hartman CA. Pilot comparison of three cardioulmonary resuscitation medication dosing strategies in overweight children. J Pediatr Pharmacol Ther 2010;15:282–9.
3. Srinivasan V, Nadkarni VM, Helfaer MA, Carey SM, Berg RA, American Heart Association National Registry of Cardiopulmonary Resuscitation I. Childhood obesity and survival after in-hospital pediatric cardiopulmonary resuscitation. Pediatrics 2010;125:e481–8.
4. Wells M, Goldstein L, Bentley A. It is time to abandon age-based emergency weight estimation in children! A failed validation of 20 different age-based formulas. Resuscitation 2017;116:73–83.
5. Luscombe M, Owens B. Weight estimation in resuscitation: is the current formula still valid? Arch Dis Child 2007;92:412–5.
6. Georgoulas V, Wells M. The PAWPER tape and the Mercy method outperform other methods of weight estimation in children in South Africa. S Afr Med J 2016;106:933–9.
7. Wells M, Coovadia A, Kramer E, Goldstein L. The PAWPER tape: a new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. Resuscitation 2013;84:227–32.
8. Wells M, Goldstein L, Bentley A. A validation study of the PAWPER XL tape: accurate estimation of both total and ideal body weight in children up to 16 years of age. Trauma Emerg Care 2017;2:1–8.
9. O'Leary F, John B, Hann L, Pegiazoglou I, McGarvey K. Estimating weight in the paediatric emergency department. Arch Dis Child 2017;102:46–52.
10. Garcia CM, Meltzer JA, Chan KN, Cunningham SJ. A validation study of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape—a new weight estimation tool. J Pediatr 2015;167:173–1771.
11. Chavez H, Peterson R, Lo K, Arel M. Weight estimation in an inner city pediatric emergency department: the effect of obesity. Am J Emerg Med 2015;33:1364–7.
12. Goldstein L, Wells M. Fat is the new normal. J Pediatr 2016;168:257.
13. Black K, Barnett P, Wolfe R, Young S. Are methods used to estimate weight in children accurate? Emerg Med 2002;14:160–5.
14. Britnell S, Taylor S, Koziol-McLain J. Emergency weight estimation lookup tables for New Zealand children aged 5-10 years. Emerg Med Australas 2016;28:558–63.
15. Summers L, Clingerman KJ, Yang X. Validation of a body condition scoring system in rhesus macaques (Macaca mulatta): assessment of body composition by using dual-energy X-ray absorptiometry. J Am Assoc Lab Anim Sci 2012;51:88–93.
16. Branstad K. Development of a Body Figure Scale and Assessment of Overweight in a Multi-Ethnic Pre-Adolescent Population. Blacksburg, VA: Virginia Polytechnic Institute and State University; 2003.
17. Tanner JM. Reliability of anthroposcopic somatotyping. Am J Phys Anthropol 1954;12:257–65.
18. Bulik CM, Wade TD, Heath AC, Martin NG, Stunkard AJ, Eaves LJ. Relating body mass index to figural stimuli: population-based normative data for Caucasians. Int J Obes Relat Metabol Disord 2001;25:1517–24.
19. Collins M. Body figure perceptions and preferences among preadolescent children. Int J Eat Disord 1991;10:199–208.
20. Truby H, Paxton SJ. The Children's Body Image Scale: reliability and use with international standards for body mass index. Br J Clin Psychol 2008;47(Pt 1):119–24.
21. Warschburger P, Kroller K. Childhood overweight and obesity: maternal perceptions of the time for engaging in child weight management. BMC Public Health 2012;12:295.
22. Lockner DW, Heyward VH, Griffin SE, Marques MB, Stolarczyk LM, Wagner DR. Cross-validation of modified fatness-specific bioelectrical impedance equations. Int J Sport Nutr 1999;9:48–59.
23. Cortes-Martinez G, Vallejo-de la Cruz N, Perez-Salgado D, Ortiz-Hernandez L. Usefulness of body silhouettes to assess the nutritional status of school-age children and adolescents in Mexico City. Bol Méd Hosp Infant México 2009;66:511–21.
24. Xie B, Avila JI, Ng BK, et al. Accurate body composition measures from whole-body silhouettes. Med Phys 2015;42:4668–77.
25. Britnell S, Koziol-McLain J. Weight estimation in paediatric resuscitation: a hefty issue in New Zealand. Emerg Med Australas 2015;27:251–6.
26. Yepes M, Viswanathan B, Bovet P, Maurer J. Validity of silhouette showcards as a measure of body size and obesity in a population in the African region: a practical research tool for general-purpose surveys. Popul Health Metrics 2015;13:35.
27. Gardner RM, Brown D. Body image assessment: a review of figural drawing scales. Person Individ Diff 2010;48:107–11.
28. Hart RP, O'Shanick GJ. Forgetting rates for verbal, pictorial, and figural stimuli. J Clin Exp Neuropsychol 1993;15:245–65.
29. Hohenhaus SM, Frush KS. Pediatric patient safety: common problems in the use of resuscitative aids for simplifying pediatric emergency care. J Emerg Nurs 2004;30:49–51.
30. Abdel-Rahman A, Jacobsen R, Watts J, et al. Comparative performance of pediatric weight estimation techniques: a human factor errors analysis. Pediatr Emerg Care 2017;33:548–52.
31. Frush K, Hohenhaus S, Luo X, Gerardi M, Wiebe RA. Evaluation of a web-based education program on reducing medication dosing error: a multicenter, randomized controlled trial. Pediatr Emerg Care 2006;22:62–70.

ARTICLE SUMMARY

1. Why is this topic important?

The PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape is one of the most accurate weight estimation systems available at the moment, but the original methodology relies on a visual assessment of a child's habitus as part of the weight estimation process. Differences in the assessment of habitus between users in different populations and between users of different skill may limit the accuracy of the system. A method of improving the accuracy of habitus assessment would be of benefit.

2. What does this study attempt to show?

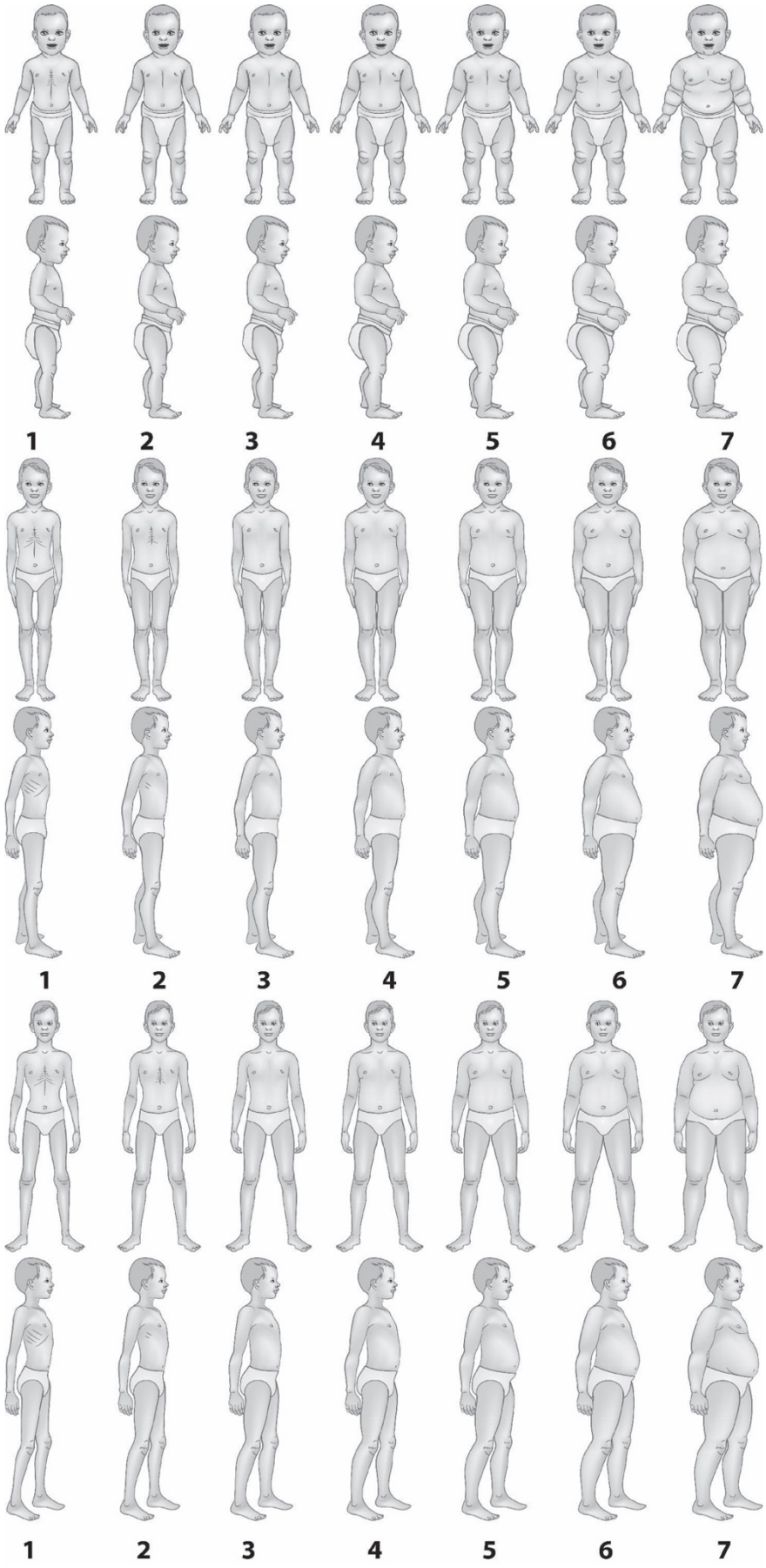
The use of figural reference images has been shown to be useful to assess habitus in other fields. We attempted to show that untrained users would be able to quantify habitus more accurately when using images than when using visual assessment or a checklist. This would result in an increased accuracy of weight estimation when using the PAWPER tape.

3. What are the key findings?

Untrained users were far more accurate at assessing habitus when using images to guide their assessment and this would have resulted in a significant improvement in their weight estimation accuracy. Users were far more satisfied with the image system when compared to the traditional system.

4. How is patient care impacted?

The use of these images in conjunction with the PAWPER tape would result in more accurate weight estimation than the original methodology. More accurate weight estimations mean more accurate drug doses and a reduction in medication errors in pediatric resuscitation. Medication errors in the emergency department are more common and probably more serious than we realize, and have a potential for causing patient harm. Any effort to reduce this is important.



Supplementary figure – Figural reference images with frontal and lateral views for three age-groups.

8.3 Epilogue

8.3.1 Major findings

Although the methodology of this study was complex, the research questions were very simple. It must be noted that the selection of children's photographs rated by the participants included children with a full range of habitus types, including extreme habitus types (very thin and morbidly obese children). This tested the methodology far more rigorously than if a sample of children that presented to any particular Emergency Department was used.

The major finding was that the PAWPER XL tape system was more accurate when habitus was assessed with the assistance of figural reference images. This was also the system that was most favoured by the participants. The improvement in accuracy was most marked in the subgroups of underweight and obese children, which suggested that the use of reference images somewhat reduced the visual biases described in the prologue.

The study also showed that nurses were not as good at evaluating habitus as doctors and paramedics. There was no convincing explanation for this. This difference in ability between individual users was interesting and it would be worthwhile in the future to evaluate the effects of training on these abilities.

The feedback from the participants on the cognitive stress associated with the different methods of habitus assessment was also unexpected. The near-universal consensus that the checklist was impractical was anticipated. Some participants felt that using gestalt to assess habitus was the quickest and least stressful method, while others felt the uncertainty caused by "guessing" was itself stressful. The overall high accuracy and high approval rating of the reference image system made it the clear choice to recommend for use with the PAWPER XL tape.

8.3.2 This study in the context of the thesis

This study showed that even minimally trained users could achieve good weight estimation results when using reference images, surpassing the accuracy of the traditional PAWPER methodology. There was, however, still a significant variation between the individual users in terms of the accuracy of habitus evaluation. The subjective elements of bias had, therefore, not been fully eliminated with the use of reference images.

There are a number of other ways of making habitus assessment more objective: anthropometric methods, other high-technological methods of quantifying habitus and automated quantification of habitus. The use of ultrasound and bioelectrical impedance measurements to quantify habitus was evaluated as part of this research but was not included in the thesis [19]. These methods showed promise but require expensive equipment and user expertise which may not be universally available. Automated habitus scoring has been accomplished in animals but has not yet been tested in humans other than in some preliminary studies on the use of mobile phone apps to assist with body composition assessment [12, 20]. Anthropometric methods of assessing habitus have been well-proven and are simple, reliable and inexpensive. The major measurement used for this purpose is mid-arm circumference, although other measurements such as waist circumference have also been shown to be useful to define habitus [21].

One of the other themes that emerged during this research was the need for adequate training and ongoing practice in weight estimation to ensure that healthcare providers achieve the maximum possible weight estimation accuracy. The large differences in individual participants' ability in habitus assessment in this study reinforced this fact. This same disparity in ability was also found in the final study, reported in Chapter 10. A strong recommendation is that training in the use of weight estimation systems is essential and this training should be a routine part of paediatric emergency simulation training.

8.3.3 Link to the next article

Mid-arm circumference has a well-established track record of use in weight estimation systems, most notably the Cattermole MAC formulas, the Wozniak method and the Mercy method [22-24]. In the study covered in the next chapter, we set out to develop a method of using mid-arm circumference to categorise habitus. This would completely eliminate the need for a visual assessment of habitus and could reduce the subjectivity of habitus scoring. It would also allow the system to be standardised between different users and different population groups, as long as users could measure mid-arm circumference accurately.

This was the next step in the process of maximising the accuracy of the PAWPER system. Since the PAWPER system has always been specifically intended for emergency use, one of the constraints in the development methodology was to ensure that no calculations would be required at any stage of the weight estimation process. The outcomes of this study, with the resultant PAWPER XL-MAC methodology are explained in the next chapter.

8.3.4 References

[1] Garcia CM, Meltzer JA, Chan KN, Cunningham SJ. A validation study of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape-a new weight estimation tool. Journal of Pediatrics. 2015;167(1):173-177.e1.

[2] O'Leary F, John-Denny B, McGarvey K, Hann A, Pegiazoglou I, Peat J. Estimating the weight of ethnically diverse children attending an Australian emergency department: a prospective, blinded, comparison of age-based and length-based tools including Mercy, PAWPER and Broselow. Archives of Disease in Childhood. 2017;102(1):46-52.

[3] Twarog JP, Politis MD, Woods EL, Daniel LM, Sonneville KR. Is obesity becoming the new normal? Age, gender and racial/ethnic differences in parental misperception of obesity as being 'About the Right Weight'. International Journal of Obesity (London). 2016;40(7):1051-1055.

[4] Goldstein L, Wells M. Fat is the new normal. Journal of Pediatrics. 2016;168:257.

[5] Cornelissen KK, Gledhill LJ, Cornelissen PL, Tovee MJ. Visual biases in judging body weight. British Journal of Health Psychology. 2016;21(3):555-69.

[6] Robinson E. Overweight but unseen: a review of the underestimation of weight status and a visual normalization theory. Obesity Reviews. 2017;18(10):1200-1209.

[7] Muttarak R. Normalization of plus size and the danger of unseen overweight and obesity in England. Obesity (Silver Spring). 2018;26(7):1125-1129.

[8] Gardner RM, Brown D. Body image assessment: A review of figural drawing scales. Personality and Individual Differences. 2010;48:107-111.

[9] Truby H, Paxton SJ. The Children's Body Image Scale: reliability and use with international standards for body mass index. British Journal of Clinical Psychology. 2008;47(Pt 1):119-24.

[10] Stunkard A. Old and new scales for the assessment of body image. Perceptual and Motor Skills. 2000;90(3 Pt 1):930.

[11] Yepes M, Viswanathan B, Bovet P, Maurer J. Validity of silhouette showcards as a measure of body size and obesity in a population in the African region: A

practical research tool for general-purpose surveys. Population Health Metrics. 2015;13:35.

[12] Xie B, Avila JI, Ng BK, Fan B, Loo V, Gilsanz V, et al. Accurate body composition measures from whole-body silhouettes. Medical Physics. 2015;42(8):4668-77.

[13] Cortes-Martinez G, Vallejo-de la Cruz N, Perez-Salgado D, Ortiz-Hernandez L. Usefulness of body silhouettes to assess the nutritional status of school-age children and adolescents in Mexico City. Boletín Médico del Hospital Infantil de México. 2009;66:511-521.

[14] Vasseur E, Gibbons J, Rushen J, de Passille AM. Development and implementation of a training program to ensure high repeatability of body condition scoring of dairy cows. J Dairy Sci. 2013;96(7):4725-37.

[15] Clingerman KJ, Summers L. Validation of a body condition scoring system in rhesus macaques (Macaca mulatta): inter- and intrarater variability. Journal of the American Association of Laboratory Animal Sciience. 2012;51(1):31-6.

[16] Vieira A, Brandao S, Monteiro A, Ajuda I, Stilwell G. Development and validation of a visual body condition scoring system for dairy goats with picture-based training. Journal of Dairy Science. 2015;98(9):6597-608.

[17] Georgoulas V, Wells M. The PAWPER tape and the Mercy Method outperform other methods of weight estimation in children in South Africa. South African Medical Journal. 2016;106(9):933-939.

[18] Truby H, Paxton SJ. Development of the Children's Body Image Scale. British Journal of Clinical Psychology. 2002;41(Pt 2):185-203.

[19] Wells M, Goldstein L, Bentley A. High-tech adjuncts to emergency weight estimation: Point-of-care ultrasound and point-of-care bioelectrical impedance measurements can increase the accuracy of length-based weight estimation in children. Trauma and Emergency Care. 2017;2(3):1-8.

[20] Bercovich A, Edan Y, Alchanatis V, Moallem U, Parmet Y, Honig H, et al. Development of an automatic cow body condition scoring using body shape signature and Fourier descriptors. J Dairy Sci. 2013;96(12):8047-59.

[21] Reilly JJ. Mid-upper arm circumference (MUAC): new applications for an old measure. Archives of Disease in Childhood. 2017;102(1):1-2.

[22] Cattermole GN, Leung PY, Mak PS, Graham CA, Rainer TH. Mid-arm circumference can be used to estimate children's weights. Resuscitation. 2010;81(9):1105-10.

[23] Whitfield KC, Wozniak R, Pradinuk M, Karakochuk CD, Anabwani G, Daly Z, et al. Anthropometric measures are simple and accurate paediatric weight-prediction proxies in resource-poor settings with a high HIV prevalence. Archives of Disease in Childhood. 2017;102(1):10-16.

[24] Abdel-Rahman SM, Ridge AL. An improved pediatric weight estimation strategy. Open Medical Devices Journal. 2012;4:87-97.

Chapter 9

Development and validation of the PAWPER XL-MAC system

9.1 Prologue

9.1.1 Background

Mid-arm circumference has been used for many years in the assessment of malnutrition in low- and middle-income countries and has been shown to be strongly associated with TBW in both children and adults [1, 2]. It has also been very successfully used as a surrogate for habitus as one component of the very accurate Mercy method [3]. While other anthropometric measurements, such as waist circumference, have been incorporated into weight estimation methodologies, mid-arm circumference remains one of the quickest, easiest and most reliable measurements to acquire [4, 5].

Since mid-arm circumference has a proven, strong association with habitus in children, adolescents and adults, it was likely that it could be used to create habitus score categories within each length segment of the PAWPER XL tape. One of the pre-conditions was that the model could not require any calculations to obtain the weight estimate (such as the Mercy method uses). For this reason, we used mid-arm circumference ranges within each length segment to define the habitus score categories and the predicted weight for each habitus score.

The use of this more objective, standardised method of habitus assessment could allow the tape to have greater uniformity of accuracy across different populations. Since other methods of habitus score assessment for the PAWPER XL tape are more subjective and dependent on experience, this new evolution of the methodology could potentially eliminate that problem.

Although the term “mid-upper arm circumference” (MUAC) is also used throughout the nutritional literature, we preferred the term ‘mid-arm circumference’ (MAC) given that it is more anatomically correct and more commonly used in the weight estimation

literature. Most anthropometric organisations simply use the term “arm circumference”, but this is not used in the medical or nutritional fields.

9.1.2 Purpose of the study

The purpose of this study was firstly to establish whether mid-arm circumference ranges could, in fact, be used in a model to determine habitus score. The second purpose was to determine whether this model could be used to estimate weight accurately in a variety of different populations, both from high-income and low- and middle-income countries.

In the broader context of this thesis, the purpose of this study would be to take the PAWPER system as far as possible in addressing the limitations identified by other studies. The specific purpose of this study was thus to explore the more objective methods of categorising habitus without needing to use a visual estimation, with its vulnerability to bias. Although there would clearly also be limitations to using anthropometric measurements, including issues relating to usability (ease-of-use and practicality in uncooperative patients), the concept needed to be explored.

This study addressed Objective 5 (to develop and validate a new, standardised and more objective system of habitus assessment for the new PAWPER XL-MAC system using measurements of mid-arm circumference).

9.2 Published paper

Wells M, Goldstein L, Bentley A. Development and validation of a method to estimate bodyweight in critically ill children using length and mid-arm circumference measurements - the PAWPER XL-MAC system. South African Medical Journal. 2017;107(11):1015-1021.

The PDF of the published article is included. The online-only supplementary material is appended after the PDF. The tables or figures have been retained in the formatting in which they are archived in the journal website.

This open-access article is distributed under Creative Commons licence CC-BY-NC 4.0.

Development and validation of a method to estimate body weight in critically ill children using length and mid-arm circumference measurements: The PAWPER XL-MAC system

M Wells, MB BCh, MSc (Med) (Emergency Medicine), FCEM (SA); **L N Goldstein**, MB BCh, MMed (Emergency Medicine), FCEM (SA), Cert Critical Care (SA); **A Bentley**, MB BCh, PhD

Division of Emergency Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Corresponding author: *M Wells (mike.wells@emergencymedicine.co.za)*

Background. Erroneous weight estimation during the management of emergency presentations in children may contribute to patient harm and poor outcomes. The PAWPER (Paediatric Advanced Weight Prediction in the Emergency Room) XL tape is an accurate length-based, habitus-modified weight estimation device, but is vulnerable to errors if subjective visual assessments of children's body habitus are incorrect or erratic.
Objective. Mid-arm circumference (MAC) has previously been used as a surrogate indicator of habitus, and the objective of this study was to determine whether MAC cut-off values could be used to predict habitus scores (HSs) to create an objective and standardised weight estimation methodology, the PAWPER XL-MAC method.
Methods. The PAWPER XL-MAC model was developed by creating MAC ranges for each HS in each weight segment of the tape. This model was validated against two samples, the National Health and Nutrition Examination Survey datasets and data from two previous PAWPER tape studies. The primary outcome measure was to achieve >70% of estimations within 10% of measured weight (PW10 >70%) and >95% within 20% of measured weight (PW20 >95%) for children aged 0 - 18 years.
Results. The PAWPER XL-MAC model achieved very high accuracy in the three validation datasets (PW10 79.2%, 79.0% and 81.9%) and a very low critical error rate (PW20 98.5%, 96.0% and 98.0%). This accuracy was maintained across all ages and in all habitus types, except for the severely obese.
Conclusions. The PAWPER XL-MAC model proved to be a very accurate, fully objective, standardised system in this study. It has the potential to be accurate across a wide variety of populations, even when used by those not experienced in visual assessment of habitus.

S Afr Med J 2017;107(11):1015-1021. DOI:10.7196/SAMJ.2017.v107i11.12505

It is important to predict the weight of children who require resuscitation, or any form of emergency medical care, accurately in order to provide appropriate doses of potentially life-saving medications.[1] Estimated weights have been shown to be extremely inaccurate in many settings, however, with some studies in underweight and obese populations having shown weight estimations to be inaccurate in up to 85% of children.[2-5] These are dangerous results that would inevitably lead to many critical medication errors. Even in populations without extreme prevalences of underweight or obesity, many commonly used methods, such as age-based formulae, are frequently inaccurate in more than half the children in whom they are used.[6] To continue to use such methods, which are known to be inaccurate, cannot be considered good medical practice.[7]

A variety of systems have been developed to improve the accuracy of weight prediction, but all have their limitations. The most accurate current weight estimation systems are the dual length- and habitus-based systems, such as the Mercy method and the PAWPER (Paediatric Advanced Weight Prediction in the Emergency Room) tape.[8,9] For these systems that rely on length to predict weight, it is important that habitus is quantified accurately in order to obtain a correct weight estimation in all children, including those who are underweight or obese.

The PAWPER XL tape (the next generation of the original PAWPER tape) was specifically designed to be used in emergencies and makes use of measurement of length as well as assessment of body habitus, either visually or with the assistance of figural reference images, to generate a rapid, accurate, calculation-free estimation of weight.[10,11] However, recent studies on the original PAWPER tape in very obese populations, and with novice users, failed to show the degree of accuracy reached in studies in less obese populations.[8,10,12-14] An important contributor to the reduced accuracy of the PAWPER tape was inaccurate assessment of body habitus in these obese children.[12,15]

The development of a less subjective and more standardised method of assessing habitus could therefore potentially improve the functioning of the PAWPER system in different populations and with users of different experience. Maximum objectivity can best be achieved with the use of a simple, easy-to-perform anthropometric measurement such as mid-arm circumference (MAC), since MAC has been shown to have a strong association with body weight, as well as body habitus, in children and adolescents.[9,16-18]

An additional issue is that, depending on the drugs to be used, there are situations where estimations of both total body weight (TBW) and ideal body weight (IBW) may be required to optimise drug administration. Obese children, in particular, could be at risk of toxicity or reduced therapeutic effect if drug doses are not corrected for body composition. In general, obese children should have lipophilic drugs (e.g. amiodarone) scaled to TBW and hydrophilic drugs scaled to IBW, even during emergency care.[19] A system that

could provide simultaneous, accurate estimations of both TBW and IBW would be valuable.

Objective

This study was an attempt to use MAC as a rapid, objective tool to predict habitus and thus, in association with measured length, predict TBW (and IBW): the PAWPER XL-MAC system.

Methods

The steps followed in developing, calibrating and validating the model are shown in Fig. 1, along with an image of the tape developed to make these data practical during emergency care.

Development of the preliminary PAWPER XL-MAC model

The original PAWPER XL tape makes use of a visual assessment of habitus to assign a habitus score (HS) (HS1 - HS7, with HS1 representing an underweight child, HS3 an average child and HS7 a severely obese child), with the predicted weights associated with each HS based on weight-for-length growth chart centiles (see Fig. 1 for details). To make use of measurements of MAC to predict HS (and TBW), rather than a visual assessment of habitus, measurement ranges of MAC measurements needed to be created for each of the seven HSs in each segment of the tape. To do this, MAC-for-length centiles were created from growth chart data downloaded from the Centers for Disease Control (CDC) website.[20] These centiles were then matched with the weight-for-length centiles used to define each HS on the original PAWPER XL tape. Finally, the centiles were used to create the MAC measurement ranges used to define the HSs.

Once this theoretical preliminary model had been constructed, it was evaluated for accuracy and the MAC measurement ranges were fine-tuned where needed.

Calibration and refinement of the PAWPER XL-MAC model

The National Health and Nutrition Examination Survey (NHANES) datasets A - G (seven datasets from the 1999 - 2000 to 2011 - 2012 surveys) were downloaded from the CDC website.[21] The demographic and anthropometric data for all children ≤18 years of age were extracted. The specific variables retained included age, height or recumbent length, TBW, body mass index (BMI) and MAC. Cases with missing or incomplete data were excluded. The 2007 - 2008 and 2009 - 2010 surveys (datasets E and F) were pooled to be used in the first round of calibration. The other datasets were used for the final validation of the refined model. BMI-for-age *z*-scores were calculated using World Health Organization reference data for children aged <24 months and CDC reference data for children from 2 to 18 years of age.

Weight estimates were generated by the preliminary model using the measurements of length and MAC from the downloaded datasets. These weight estimates were then compared against the actual measured weight for each child. The accuracy of the estimations was evaluated in the entire sample as well as in each of the 34 weight segments of the PAWPER XL-MAC tape model.

The cut-off MAC values were adjusted (calibrated) until an acceptable degree of accuracy was obtained in each segment (see below for acceptable outcome measures). The outcome data before and after calibration are shown in Table I. The final model, the details of which are shown in Table II, was then subjected to a validation assessment.

Validation of the final PAWPER XL-MAC model

The final model was validated in three samples: the pooled unused NHANES data (datasets A - D and G) and data from two previous

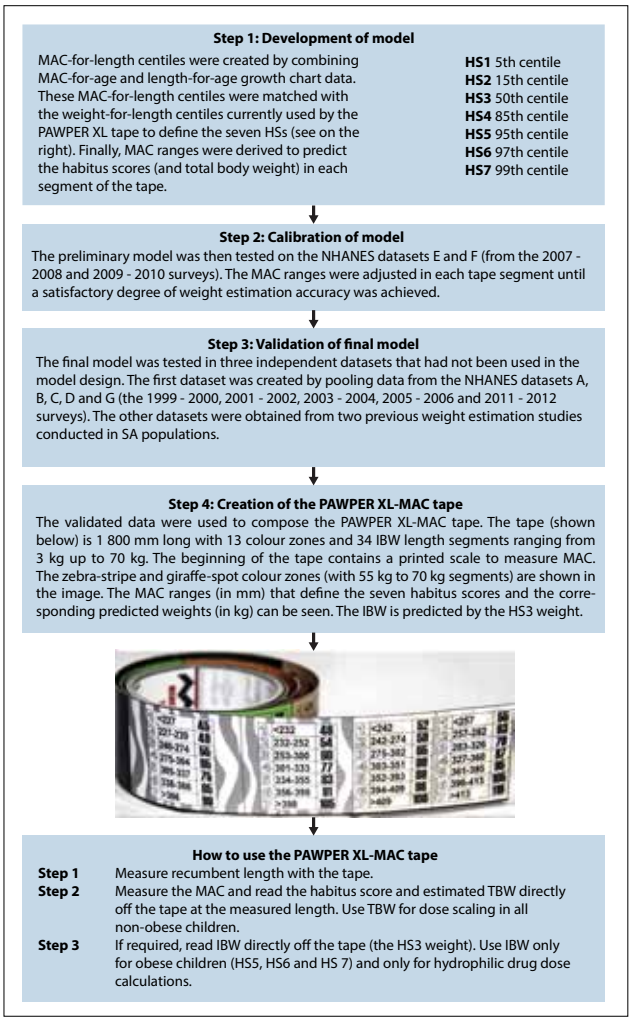


Fig. 1. The methodology followed in the study. The process of creating the initial model, calibrating this model and then validating the final model in three independent samples is illustrated. The figure also shows how the validated data were converted into the full-length PAWPER XL-MAC tape, which could be used in clinical practice. Both TBW (adjusted for MAC-determined habitus) and IBW can be read directly off the tape. (MAC = mid-arm circumference; HS = habitus score; NHANES = National Health and Nutrition Examination Survey; SA = South African; IBW = ideal body weight; TBW = total body weight.)

PAWPER tape weight estimation studies in which MAC data were captured.[10,11] These two studies were prospective, cross-sectional studies conducted in two hospitals in Johannesburg, South Africa (SA). One of the hospitals serves a community of middle to upper socioeconomic status and the other a community of mostly low socioeconomic status. Study 1 enrolled 332 children from July 2014 to December 2014 and study 2 enrolled 300 children from August 2014 to January 2015. As in the calibration step, weight estimates were generated by the final model using the measurements of length and MAC from the validation datasets. These weight estimates were then compared against the measured weight for each child in the entire sample as well as in each of the individual segments.

Statistical analysis

The performance of the PAWPER XL-MAC model in predicting measured weight was evaluated using three major statistical measures: the mean percentage error (MPE) represented the

Table 1. Accuracy, bias and precision of the PAWPER XL-MAC model during development, before and after calibration and validation*

Dataset	*N*	PW10	PW20	MPE (LLOA, ULOA)
Accuracy of primary derivation model (before calibration)				
NHANES datasets E, F	4 664	53.2	86.2	−9.6 (−28.7, 9.6)
Accuracy of model after calibration				
NHANES datasets E, F	4 664	84.9	98.7	−2.1 (−16.8, 12.6)
Accuracy of final model in validation samples				
NHANES datasets A, B, C, D, G	13 134	81.9	98.0	−0.9 (−17.5, 15.7)
Study 1	332	79.2	98.5	1.5 (−14.5, 17.5)
Study 2	300	79.0	96.0	5.0 (−11.0, 21.0)
Original PAWPER tape methodology performance in SA validation samples				
Study 1	332	83.4	98.5	1.1 (−13.9, 16.1)
Study 2	300	88.0	97.7	1.4 (−12.5, 15.3)

MAC = mid-arm circumference; PW10 and PW20 = percentage of weight estimations falling within 10% and 20% of actual weight; MPE = mean percentage error; LLOA and ULOA = lower and upper limits of agreement; NHANES = National Health and Nutrition Examination Survey; SA = South Africa/n.
*The performance of the primary model before calibration is shown in the top section. PW10 and PW20 provide a quantification of accuracy. The MPE indicates the bias of the model, with a negative value denoting a bias to underestimation of weight. LLOA and ULOA provide an indication of precision: the ideal model should have 95% of estimations within ±20% of actual weight. The performance of the adjusted model (after calibration) in the derivation sample (NHANES datasets E and F from the 2007 - 2008 and 2009 - 2010 surveys) is shown in the second section. The performance of the final model in three new validation samples is shown in the bottom two sections. The NHANES dataset was derived from children in the USA (NHANES datasets A, B, C, D and G from the 1999 - 2000, 2001 - 2002, 2003 - 2004, 2005 - 2006 and 2011 - 2012 surveys), and the study 1 and study 2 datasets were derived from children in Johannesburg, SA. The final model achieved the acceptable outcome criteria of PW10 >70% and PW20 >95% in all three of the validation datasets. The validation of the final model was also performed independently in each of the 34 length segments of the tape, with the acceptable outcome criteria being achieved in every individual segment. The bottom panel shows the performance data of the original PAWPER methodology from the SA validation datasets, for comparison with the new model.

estimation bias; the 95% limits of agreement of the MPE quantified the estimation precision; and the percentage of weight estimations that fell within 10% (PW10) and 20% (PW20) of measured weight denoted overall accuracy. Subgroup analyses were performed on the pooled data for children of different weight status (underweight to severely obese).

All data were analysed using Microsoft Excel version 15.38, 2016 (Microsoft, USA) and Stata Statistical Software release 14, 2015 (StataCorp, USA).

Outcome measures

The primary outcome measure was the accuracy of the PAWPER XL-MAC model's estimation of weight when compared with measured weight. The secondary outcome measure was comparing the performance of the model with the accuracy of the original PAWPER tape method in the SA validation samples.

No previous weight estimation study has recommended an appropriate benchmark by which to characterise acceptable accuracy for a weight estimation system. It is, however, statistically important to have *a priori* criteria by which to judge the functioning of any methodology. We used a PW10 >70% together with a PW20 >95% to define acceptable accuracy of weight estimation. This was modified from the criteria proposed in an Australian biostatistical research report on the accuracy of the Broselow tape.[22] This target also matches the level of accuracy generally achieved by the most accurate existing weight estimation systems. A weight estimation error of >20% was considered critical because of the resultant high risk of medication error.

Results

Characteristics of study participants

The demographic characteristics of the children included in the calibration and validation studies are shown in Table III. The SA samples included a greater proportion of younger and underweight children than the NHANES dataset, although the study 1 sample contained a wide variety of body types with a high prevalence of both underweight and obese children. The study 2 sample was from an impoverished community with a very high prevalence of underweight children. Together the validation datasets provided children with a broad range of ages and body types in which to rigorously evaluate the model.

Validation of the PAWPER XL-MAC method

Primary outcome measures – achieving acceptable accuracy

As can be seen in Table I, the PAWPER XL-MAC method exceeded the acceptable outcome criteria (PW10 >70% and PW20 >95%) in each of the three validation samples overall, as well as in every segment-by-segment analysis. The best performance was in the 17 kg segment, with a PW10 of 95.7% and a PW20 of 99.5%, and the poorest performance was in the 60 kg segment, with a PW10 of 73.3% and a PW20 of 97.1%.

Bland-Altman plots illustrating the performance of the final model in the pooled validation samples, as well as for subgroups representing extremes of habitus, are shown in Fig. 2.

Subgroup analysis by extremes of body habitus (Table IV) showed that the primary outcome measures were fully met in five of six categories. Only in severely obese children was the accuracy poor (PW10 43.2%) and the critical error rate higher than the primary outcome measures permitted (PW20 80.6%).

Secondary outcome measures – comparison with original PAWPER methodology

The performance of the PAWPER XL-MAC method in the study 1 population, which included extremes of body habitus, was similar to the original results with the regular PAWPER XL methodology (see bottom of Table I). In the study 2 population, the PAWPER XL-MAC method performed acceptably well, but with a higher bias to overestimation of weight than with visual habitus assessment. The model was not as accurate as the original method (PW10 comparisons, odds ratio 2.2 (confidence interval 1.4 - 3.4); p=0.0012 (Fisher's exact test)). There was no difference in the critical error rate (PW20), however.

Discussion

Main findings

The PAWPER XL-MAC model satisfied the primary outcome measures by achieving a PW10 in >70% of children, and a critical error rate of <5% (PW20 >95%). The findings were consistent across all lengths and in all habitus types except for severely obese children, who accounted for just over 4% of the validation sample. The model was also accurate in two smaller SA validation samples, achieving a reasonably similar accuracy to the conventional PAWPER methodology. The diverse demographic and anthropometric charac-

Table 2. The detailed data describing the full PAWPER XL-MAC tape model*

		HS1		HS2			HS3			HS4			HS5			HS6			HS7	
Recumbent length (cm)	IBW (kg)	TBW (kg)	MAC upper limit (cm)	TBW (kg)	MAC lower limit (cm)	MAC upper limit (cm)	TBW (kg)	MAC lower limit (cm)	MAC upper limit (cm)	TBW (kg)	MAC lower limit (cm)	MAC upper limit (cm)	TBW (kg)	MAC lower limit (cm)	MAC upper limit (cm)	TBW (kg)	MAC lower limit (cm)	MAC upper limit (cm)	TBW (kg)	MAC lower limit (cm)
43.0 - 49.3	3	2.5	<9.4	2.8	9.4	10.3	3	10.4	11.5	3.2	11.6	12.7	3.5	12.8	13.9	4	14.0	15.1	4.5	15.2
49.4 - 54.8	4	3.5	<9.6	3.8	9.6	10.4	4	10.5	11.7	4.5	11.8	12.9	5	13.0	14.1	6	14.2	15.4	7	15.5
54.9 - 59.2	5	4	<11.0	4.5	11.0	11.4	5	11.5	12.4	5.5	12.5	13.4	6	13.5	15.4	7	15.5	16.4	8	16.5
59.3 - 62.9	6	5	<11.5	5.5	11.5	11.9	6	12.0	13.1	6.5	13.2	14.2	7	14.3	15.2	8	15.3	16.4	9	16.5
63.0 - 66.4	7	6	<11.5	6.5	11.5	12.9	7	13.0	14.4	7.5	14.5	15.3	8	15.4	16.4	9	16.5	17.4	10	17.5
66.5 - 70.1	8	7	<12.5	7.5	12.5	13.6	8	13.7	14.4	8.5	14.5	15.4	9.5	15.5	16.9	10	17.0	18.4	11	18.5
70.2 - 73.9	9	7.5	<12.6	8.5	12.6	14.1	9	14.2	15.4	10	15.5	16.2	11	16.3	18.1	12	18.2	19.9	13	20.0
74.0 - 78.2	10	8.5	<12.7	9.5	12.7	14.4	10	14.5	15.4	11	15.5	17.0	12	17.1	18.6	13	18.7	20.4	14	20.5
78.3 - 83.1	11	9	<12.7	10	12.7	14.6	11	14.7	15.4	12	15.5	17.0	13	17.1	18.9	14	19.0	20.4	15	20.5
83.2 - 88.0	12	10	<13.0	11	13.0	14.8	12	14.9	15.7	13	15.8	17.1	14	17.2	18.9	16	19.0	20.4	18	20.5
88.1 - 92.6	13	11	<13.5	12	13.5	14.8	13	14.9	15.8	14	15.9	17.2	15	17.3	19.0	17	19.1	20.9	19	21.0
92.7 - 96.8	14	12	<14.2	13	14.2	15.1	14	15.2	16.5	15	16.6	17.6	17	17.7	19.3	19	19.4	20.9	21	21.0
96.9 - 100.6	15	13	<14.4	14	14.4	15.6	15	15.7	16.5	16	16.6	17.6	18	17.7	19.4	20	19.5	21.4	22	21.5
100.7 - 103.9	16	14	<14.8	15	14.8	15.7	16	15.8	17.0	17	17.1	17.9	19	18.0	19.7	21	19.8	22.4	23	22.5
104.0 - 107.0	17	15	<15.3	16	15.3	15.9	17	16.0	17.4	18	17.5	18.0	20	18.1	19.9	22	20.0	22.4	24	22.5
107.1 - 110.0	18	16	<15.6	17	15.6	16.6	18	16.7	17.8	19	17.9	18.6	21	18.7	21.9	24	22.0	23.4	26	23.5
110.1 - 113.2	19	16	<15.6	18	15.6	16.6	19	16.7	17.8	21	17.9	20.2	23	20.3	22.5	25	22.6	24.4	28	24.5
113.3 - 116.5	20	17	<15.6	19	15.6	17.5	20	17.6	18.4	22	18.5	20.2	24	20.3	22.9	26	23.0	24.4	29	24.5
116.6 - 120.6	22	20	<16.3	21	16.3	17.7	22	17.8	19.1	24	19.2	21.2	27	21.3	23.1	29	23.2	24.8	32	24.9
120.7 - 125.4	24	21	<16.5	23	16.5	18.2	24	18.3	19.4	26	19.5	21.8	29	21.9	23.4	32	23.5	24.8	35	24.9
125.5 - 129.6	26	23	<17.5	25	17.5	18.5	26	18.6	20.4	28	20.5	22.0	31	22.1	24.4	34	24.5	25.4	37	25.5
129.7 - 133.3	28	24	<17.8	26	17.8	18.7	28	18.8	20.7	30	20.8	22.3	33	22.4	25.6	36	25.7	26.6	40	26.7
133.4 - 136.6	30	26	<17.9	28	17.9	19.4	30	19.5	22.0	33	22.1	22.9	36	23.0	25.6	40	25.7	26.6	44	26.7
136.7 - 139.8	32	28	<19.2	30	19.2	20.4	32	20.5	22.3	35	22.4	23.9	39	24.0	25.6	42	25.7	26.9	47	27.0
139.9 - 143.2	34	29	<19.2	32	19.2	20.7	34	20.8	23.2	38	23.3	24.6	42	24.7	26.5	46	26.6	28.2	50	28.3
143.3 - 146.5	36	29	<19.5	33	19.5	21.0	36	21.1	23.3	39	23.4	24.6	43	24.7	26.9	48	27.0	29.9	52	30.0
146.6 - 149.8	38	33	<20.2	36	20.2	21.6	38	21.7	23.3	41	23.4	24.9	45	25.0	26.9	50	27.0	28.6	58	28.7
149.9 - 153.1	40	36	<20.8	39	20.8	22.5	40	22.6	24.1	48	24.2	26.1	52	26.2	27.8	57	27.9	29.9	65	30.0
153.2 - 158.0	45	38	<21.1	43	21.1	23.4	45	23.5	25.5	55	25.6	28.1	64	28.2	31.5	72	31.6	35.1	84	35.2
158.1 - 165.0	50	40	<22.0	45	22.0	23.8	50	23.9	26.1	60	26.2	30.1	72	30.2	32.7	79	32.8	35.3	87	35.4
165.1 - 170.0	55	45	<22.7	49	22.7	23.9	55	24.0	27.4	65	27.5	30.4	75	30.5	33.7	85	33.8	36.6	99	36.7
170.1 - 174.0	60	46	<23.2	54	23.2	25.2	60	25.3	30.0	77	30.1	33.3	83	33.4	35.5	91	35.6	39.8	105	39.9
174.1 - 177.0	65	52	<24.2	59	24.2	27.4	65	27.5	30.2	80	30.3	35.1	89	35.2	39.3	98	39.4	40.9	108	41.0
177.1 - 180.0	70	55	<25.7	63	25.7	28.2	70	28.3	32.6	87	32.7	36.0	95	36.1	39.5	105	39.6	41.3	116	41.4

MAC = mid-arm circumference; HS = habitus score (HS1 - HS7, HS1 representing an underweight child, HS3 an average child and HS7 a severely obese child); IBW = ideal body weight; TBW = total body weight.
*The child's recumbent length determines into which length segment they would fall. The HS3 weight for each segment reflects the IBW for a child of that length (i.e. a child with a measured recumbent length of between 43.0 and 49.3 cm would be predicted to have an IBW of 3 kg). The upper and lower limits of MAC measurements that were used to define each HS within each length segment are shown. The final adjusted or predicted TBW for children of below or above average weight-for-length are shown for each HS in each segment (i.e. a child with a recumbent length of 137.0 cm and an MAC of 26.0 cm would be predicted to have an IBW of 32 kg but a TBW of 42 kg – an obese child). A full-length tape was designed from these data so that both TBW and IBW could be rapidly determined during emergency care.

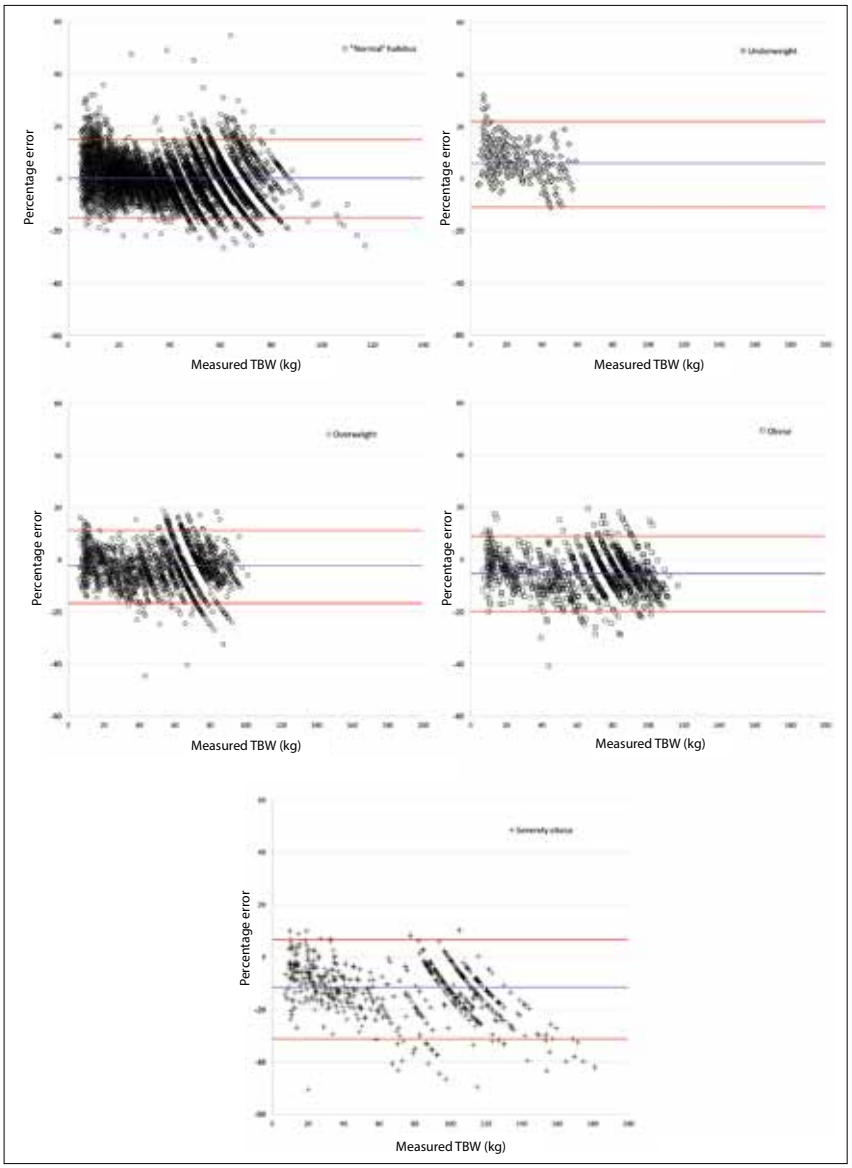


Fig. 2. Bland-Altman plots of percentage error against actual weight for the pooled validation datasets for children with 'average' and extremes of habitus. The blue lines represent the mean percentage error (trueness or bias) and the red lines represent the 95% limits of agreement (i.e. the range in which 95% of the estimates fall). (TBW = total body weight.)

teristics of the validation samples suggest that this methodology may be equally accurate in other similar populations.

The strength of the model derived in this study was its consistency and accuracy in >95% of the sample. The progression in expectations of weight estimation systems from those that achieve accuracy in <25% of children to the newer methodologies in which >70% receive an accurate estimation and <5% have a critical error, is encouraging.[9,23] It is possible that severely obese children will remain a challenge with regard to accurate weight estimation and appropriate drug dose determination, however, especially children with BMI-for-age *z*-scores well above 5.

Visual assessment of HS

Although the PAWPER tape has been shown to outperform other methods of weight estimation in children in some studies, erratic subjective assessment of body habitus by users has resulted in less-than-desirable accuracy in other studies and populations.[12-14] Accurate and repeatable anthroposcopic assessment of habitus is possible by experienced practitioners, but assessment has been shown to be less reliable when performed by novices.[24,25] Furthermore, healthcare professionals, and parents, have more difficulty in recognising overweight and obese children than underweight children, which would lead to an underestimation of HS when using the regular PAWPER XL tape methodology.[15,26] Figural reference images, validated against BMI, have been used successfully to reduce the subjectivity of anthroposcopic assessment of habitus, mostly in body image research.[27] Research on the use of figural reference images with the PAWPER XL tape has shown good results, with equal accuracy when compared with experts in visual gestalt assessment.[10] Although the use of figural reference images has been shown to improve visual assessment of habitus, it remains somewhat subjective.[10,28] A fully objective method could therefore have an important role, even if used as a secondary or confirmatory technique for the PAWPER XL tape.

In this study, the accuracy achieved by the PAWPER XL-MAC model was comparable to that obtained by visual assessment of habitus. Although the accuracy of the original PAWPER tape using visual assessment of habitus was statistically significantly better than the new PAWPER XL-MAC model in the study 2 population, the new model still performed very well. The evidence suggests that extraordinarily high levels of accuracy are achievable in underweight and normal-weight children using visual assessment, especially with skilled users, but that this is more difficult in obese children.[15,29] This disparity makes gestalt visual assessment vulnerable to error and may potentially increase cognitive stress, especially in inexperienced users. An objective, standardised, easy-to-use method is of value.

MAC and body habitus

MAC has a strong evidence base supporting its value as a surrogate for body habitus and is a reliable measurement with high inter-user agreement.[30] When used as a single variable, MAC is more accurate than age in predicting weight.[31] It is the habitus-surrogate measurement used with the Mercy method, which has been shown to be very accurate in a wide range of underweight and obese populations.[9,32] The accuracy of the PAWPER XL-MAC model in this study was similar to, or better than, the Mercy method outcomes reported previously (PW10 63.9 - 80.1% range), but far better than MAC used as a single variable (PW10 15.5 - 44.2%).[9,32] The use of MAC in the PAWPER XL-MAC model to predict HS was accurate and objective, fulfilling the main objectives of this study. This model does not change how IBW may be rapidly estimated in obese children (the HS3 weight is used, which is printed on the tape), if required. With weight estimation systems designed for use during emergency care, however, how the systems work may be as important as how accurate they are.[33]

Table 3. Demographic characteristics of children in the derivation and validation datasets*

	Calibration dataset, NHANES: USA pooled data (datasets E, F)	Validation datasets		
		NHANES: USA pooled data, datasets A, B, C, D, G	Study 1: SA, low- and middle-income groups	Study 2: SA, low-income group
N	4 664	13 134	332	300
Age (yr), median (IQR)	10.0 (4.0 - 14.0)	10.0 (4.0 - 14.0)	7.2 (4.5 - 9.3)	4.3 (2.3 - 7.0)
Sex male, *n* (%)	2 326 (49.9)	6 483 (49.4)	154 (46.7)	152 (50.7)
Length (cm), median (IQR)	145.2 (104.0 - 161.3)	142.4 (104.3 - 162)	122.0 (106.0 - 137.0)	104.0 (87.0 - 122.0)
Weight (kg), median (IQR)	40.5 (17.2 - 59.3)	38.1 (17.3 - 58.6)	23.4 (17.6 - 33.8)	16.2 (12.3 - 22.2)
BMI (kg/m^2), median (IQR)	18.7 (16.4 - 22.8)	18.7 (16.5 - 22.5)	16.7 (15.2 - 18.8)	15.2 (14.2 - 16.4)
z-score, median (IQR)	0.6 (–0.2 - 1.5)	0.5 (–0.3 - 1.4)	0.4 (–0.5 - 1.1)	–0.7 (–1.7 - 0.2)
Underweight, *n* (%)	77 (2.0)	348 (2.6)	23 (7.0)	43 (18.3)
Overweight, *n* (%)	522 (13.6)	1 686 (12.8)	38 (11.5)	12 (5.1)
Obese, *n* (%)	368 (9.6)	1 074 (8.2)	20 (6.1)	4 (1.7)
Severely obese, *n* (%)	161 (4.2)	572 (4.4)	15 (4.5)	1 (0.4)

NHANES = National Health and Nutrition Examination Survey; SA = South Africa; BMI = body mass index.
*There were substantial differences between the ages and weight status of the three validation samples. Study 2, which comprised mainly children from a low-income group, had a much higher number of underweight children than the other datasets. Study 1, which included children from both low- and middle-income groups, had a greater variety of body types than the other two datasets. The three datasets provided a reasonable variation of age groups and habitus types in which to assess the model performance.

Table 4. Subgroup analysis of the PAWPER XL-MAC model in the pooled validation datasets, by body composition*

Description	BMI-for-age *z*-score	*n* (%)	PW10	PW20
Underweight	≤2	414 (3.0)	71.3	96.3
Thin	–1.4 - –1.9	511 (3.7)	82.2	98.4
'Average'	–1.3 - 1.3	9 419 (68.4)	85.7	99.1
Overweight	1.4 - 1.9	1 736 (12.6)	82.7	98.6
Obese	2 - 2.4	1 098 (8.0)	73.1	98.1
Severely obese	≥2.5	588 (4.3)	43.2	80.6

BMI = body mass index; PW10 and PW20 = percentage of weight estimations falling within 10% and 20% of actual weight.
*There were 13 766 total available *z*-scores for inclusion. The model failed to achieve acceptable accuracy only in the severely obese group of children – the *z*-scores in this group ranged from 2.5 to 7.9, with the children's weight ranging up to 181 kg.

Constraints on the use of weight estimation systems during resuscitation

The cognitive stresses incurred by healthcare providers during emergency care reduce their ability to perform even the simplest mental tasks.[33,34] Calculation errors are particularly common, which may make any system that requires arithmetic vulnerable when used during emergencies.[35] The PAWPER XL-MAC model was therefore developed so that MAC cut-off values (ranges) could be read rapidly off the tape to determine the HS, without requiring calculations – a cognitively neutral process. This is a more complex process than visual assessment of habitus, but also one that would not instil doubts about the accuracy of habitus assessment. Adequate training would clearly be required to use the PAWPER XL-MAC system accurately – even the most basic of weight estimation techniques (such as the Broselow tape) have been shown to be performed very poorly when used by individuals with no training.[36,37] Given the somewhat increased complexity of the PAWPER XL-MAC system, a staged or protocol-driven procedure for weight estimation and weight estimation training may be required.

Staged approach to weight estimation

The use of the PAWPER XL-MAC method would be most suitable if implemented in a staged approach to weight estimation, timing the procedures to the priorities and dictates of the emergency medical care. A visual or image-assisted PAWPER XL tape weight could be obtained immediately emergency, life-saving care is initiated, followed by a confirmatory measurement-based PAWPER XL-MAC estimation several minutes later once exact drug doses are required, prior to their administration. Weight estimation practices should also be incorporated into paediatric resuscitation courses and regular simulation training. This concept would need additional evaluation in simulation-based training.

TBW and IBW estimations

It has been speculated that the use of inaccurate or inappropriate weight estimations may worsen outcomes in obese children.[38] Since there is now a reasonable consensus on which drugs should be scaled to TBW and which to IBW in obese children, these should be applied during emergency care.[39] The PAWPER XL-MAC model could facilitate the estimation of both, in keeping with recent guidelines. How this may provide benefits for care still needs to be investigated.

Study limitations

There were relatively small numbers of infants weighing <6 kg in the calibration and validation datasets, so it is uncertain whether the accuracy of the PAWPER XL-MAC technique can be generalised to this weight group, and further confirmation will be required. Measurement of MAC in a supine child receiving emergency care by a stressed healthcare provider may not be as accurate as that performed on a seated child by an anthropometry expert. This also requires further research in real-world or simulation situations.

Conclusions

The consistent level of high accuracy achieved by the PAWPER XL-MAC method across a broad age range of children in this study

exceeded the stringent outcome measures. Weight estimates were also accurate in children of all habitus types except for the severely obese (BMI-for-age *z*-score ≥2.5), in whom accurate weight estimation still proves challenging. Critical error rates were generally very low. The PAWPER XL-MAC method achieved similar accuracy of weight estimation to the original PAWPER methodology (using gestalt visual assessment) from previous studies.

The objectiveness of the PAWPER XL-MAC technique may make it more standardised and objective for novice users and those not experienced in or skilled at anthroposcopic habitus assessment. The immediate feedback with regard to HS assessment could also assist users in learning the visual method of habitus scoring.

The validation samples comprised a very diverse group of children, with the model performing well in almost all the subgroups, supporting speculation that the model may be accurate in other populations with a similar demographic composition. However, the PAWPER XL-MAC needs prospective testing in a variety of populations to establish its true accuracy and generalisability.

Acknowledgements. None.

Author contributions. The authors all contributed to the study design, analysis and interpretation of data and drafting and writing of the article.

Funding. None.

Conflicts of interest. None.

1. Kaufmann J, Laschat M, Wappler F. Medication errors in pediatric emergencies: A systematic analysis. Dtsch Arztebl Int 2012;109(38):609-616. https://doi.org/10.3238/arztebl.2012.0609
2. Clark MC, Lewis RJ, Fleischman RJ, Ogunniyi AA, Patel DS, Donaldson RI. Accuracy of the Broselow tape in South Sudan, 'the hungriest place on earth'. Acad Emerg Med 2016;23(1):21-28. https://doi.org/10.1111/acem.12854
3. Bowen L, Zyambo M, Snell D, Kinnear J, Bould MD. Evaluation of the accuracy of common weight estimation formulae in a Zambian paediatric surgical population. Anaesthesia 2017;72(4):470-478. https://doi.org/10.1111/anae.13780
4. Marlow R, Lo D, Walton L. Accurate paediatric weight estimation by age: Mission impossible? Arch Dis Child 2011;96(Suppl 1):A1-A2. https://doi.org/10.1136/adc.2011.212563.3
5. Cattermole GN, Leung MP, So HK, Mak PS, Graham CA, Rainer TH. Age-based formulae to estimate children's weight in the emergency department. Emerg Med J 2011;28(5):390-396. https://doi.org/10.1136/emj.2009.090357
6. Wells M, Goldstein L, Bentley A. It is time to abandon age-based emergency weight estimation in children! A failed validation of 20 different age-based formulas. Resuscitation 2017;116(7):73-83. https://doi.org/10.1016/j.resuscitation.2017.05.018
7. Luscombe M, Owens B. Weight estimation in resuscitation: Is the current formula still valid? Arch Dis Child 2007;92(5):412-415. https://doi.org/10.1136/adc.2006.107284
8. Wells M, Coovadia A, Kramer E, Goldstein L. The PAWPER tape: A new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. Resuscitation 2013;84(2):227-232. https://doi.org/10.1016/j.resuscitation.2012.05.028
9. Abdel-Rahman SM, Ridge AL. An improved pediatric weight estimation strategy. Open Med Devices J 2012;4:87-97. https://doi.org/10.2174/1875181401204010087
10. Georgoulas V, Wells M. The PAWPER tape and the Mercy method outperform other methods of weight estimation in children in South Africa. S Afr Med J 2016;106(9):933-939. https://doi.org/10.7196/SAMJ.2016.v106i9.10572
11. Wells M, Goldstein L, Bentley A. A validation study of the PAWPER XL tape: Accurate estimation of both total and ideal body weight in children up to 16 years of age. Trauma Emerg Care 2017;2(4):1-8. https://www.oatext.com/pdf/TEC-2-141.pdf (accessed 27 September 2017).
12. Chavez H, Peterson R, Lo K, Arel M. Weight estimation in an inner city pediatric emergency department: The effect of obesity. Am J Emerg Med 2015;33(10):1364-1367. https://doi.org/10.1016/j.ajem.2015.07.040
13. Garcia CM, Meltzer JA, Chan KN, Cunningham SJ. A validation study of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape – a new weight estimation tool. J Pediatr 2015;167(1):173-177.e1. https://doi.org/10.1016/j.jpeds.2015.04.006
14. O'Leary F, John-Denny B, McGarvey K, Hann A, Pegiazoglou I, Peat J. Estimating the weight of ethnically diverse children attending an Australian emergency department: A prospective, blinded, comparison of age-based and length-based tools including Mercy, PAWPER and Broselow. Arch Dis Child 2017;102(1):46-52. https://doi.org/10.1136/archdischild-2016-310917
15. Goldstein L, Wells M. Fat is the new normal. J Pediatr 2016;168:257. https://doi.org/10.1016/j.jpeds.2015.09.049
16. Cattermole GN, Leung M, Mak P, So HK, Graham CA, Rainer TH. Children's weights correlate more strongly with mid-arm circumference (MAC) than with age, height or foot-length. J Emerg Med 2009;37(2):228-229. https://doi.org/10.1016/j.jemermed.2009.06.051
17. Cattermole GN, Graham CA, Rainer TH. Mid-arm circumference can be used to estimate weight of adult and adolescent patients. Emerg Med J 2016;34(4). https://doi.org/10.1136/emermed-2015-205623
18. Whitfield KC, Wozniak R, Pradinuk M, et al. Anthropometric measures are simple and accurate paediatric weight-prediction proxies in resource-poor settings with a high HIV prevalence. Arch Dis Child 2017;102(1):10-16. https://doi.org/10.1136/archdischild-2015-309645
19. Kendrick JG, Carr RR, Ensom MH. Pediatric obesity: Pharmacokinetics and implications for drug dosing. Clin Ther 2015;37(9):1897-1923. https://doi.org/10.1016/j.clinthera.2015.05.495
20. National Center for Health Statistics, Centers for Disease Control and Prevention. Anthropometric Reference Data for Children and Adults: United States, 2007 - 2010. Washington, DC: CDC/NCHS, 2012.
21. National Center for Health Statistics, Centers for Disease Control and Prevention. National Health and Nutrition Examination Survey Data. 2016. https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx (accessed 27 September 2017).
22. Stewart D. Accuracy of the Broselow tape for estimating paediatric weight in two Australian emergency departments. University of Sydney, 2009. https://ses.library.usyd.edu.au//bitstream/2123/6265/1/Declan Stewart WPP final.pdf (accessed 27 September 2017).
23. Seddon C, Lockitt L, Dhanjal S, Eisenhut M. Validation of advanced paediatric life support formulas for weight calculation in a multiethnic population. International Scholarly Research Notices. Pediatrics 2012(2012), article ID 869634. https://doi.org/10.5402/2012/869634
24. Lockner DW, Heyward VH, Griffin SE, Marques MB, Stolarczyk LM, Wagner DR. Cross-validation of modified fatness-specific bioelectrical impedance equations. Int J Sport Nutr 1999;9(1):48-59. https://doi.org/10.1123/ijsn.9.1.48
25. Black K, Barnett P, Wolfe R, Young S. Are methods used to estimate weight in children accurate? Emerg Med 2002;14(2):160-165. https://doi.org/10.1046/j.1442-2026.2002.00311.x
26. Warschburger P, Kroller K. Childhood overweight and obesity: Maternal perceptions of the time for engaging in child weight management. BMC Public Health 2012;12:295. https://doi.org/10.1186/1471-2458-12-295
27. Truby H, Paxton SJ. The Children's Body Image Scale: Reliability and use with international standards for body mass index. Br J Clin Psychol 2008;47(Pt 1):119-124.
28. Britnell S, Taylor S, Koziol-McLain J. Emergency weight estimation lookup tables for New Zealand children aged 5 - 10 years. Emerg Med Australas 2016;28(5):558-563. https://doi.org/10.1111/1742-6723.12637
29. Twarog JP, Politis MD, Woods EL, Daniel LM, Sonneville KR. Is obesity becoming the new normal? Age, gender and racial/ethnic differences in parental misperception of obesity as being 'about the right weight'. Int J Obes (Lond) 2016;40(7):1051-1055. https://doi.org/10.1038/ijo.2016.40
30. WHO Multicentre Growth Reference Study Group. Reliability of anthropometric measurements in the WHO Multicentre Growth Reference Study. Acta Paediatr 2006;450(Suppl):38-46. https://doi.org/10.1111/j.1651-2227.2006.tb02374.x
31. Cattermole GN, Leung PY, Mak PS, Graham CA, Rainer TH. Mid-arm circumference can be used to estimate children's weights. Resuscitation 2010;81(9):1105-1110. https://doi.org/10.1016/j.resuscitation.2010.05.015
32. Dicko A, Alhousseini ML, Sidibé B, Traoré M, Abdel-Rahman SM. Evaluation of the Mercy weight estimation method in Ouelessebougou, Mali. BMC Public Health 2014;14:270. https://doi.org/10.1186/1471-2458-14-270
33. Luten R, Wears RL, Broselow J, Croskerry P, Joseph MM, Frush K. Managing the unique size-related issues of pediatric resuscitation: Reducing cognitive load with resuscitation aids. Acad Emerg Med 2002;9(8):840-847. https://doi.org/10.1197/aemj.9.8.840
34. Lancaster L. Mission impossible – minimising the terror of paediatric resuscitation for staff in the ED. Australas Emerg Nurs J 2002;5(3):16-20. https://doi.org/10.1016/S1328-2743(02)80004-9
35. Pentin J, Smith J. Drug calculations: Are they safer with or without a calculator? Br J Nurs 2006;15(14):778-781. https://doi.org/10.12968/bjon.2006.15.14.21582
36. Abdel-Rahman A, Jacobsen R, Watts J, et al. Comparative performance of pediatric weight estimation techniques: A human factor errors analysis. Pediatr Emerg Care 2015;33(8):548-552. https://doi.org/10.1097/PEC.0000000000000543
37. Hohenhaus SM, Frush KS. Pediatric patient safety in resuscitation simulations: Lessons learned from studying clinician behavior. Pediatr Emerg Care 2005;21(1):71-75. https://doi.org/10.1097/01.pec.0000150992.49722.60
38. Srinivasan V, Nadkarni VM, Helfaer MA, Carey SM, Berg RA, American Heart Association National Registry of Cardiopulmonary Resuscitation I. Childhood obesity and survival after in-hospital pediatric cardiopulmonary resuscitation. Pediatrics 2010;125(3):e481-e488. https://doi.org/10.1542/peds.2009-1324
39. Ross EL, Heizer J, Mixon MA, et al. Development of recommendations for dosing of commonly prescribed medications in critically ill obese children. Am J Health Syst Pharm 2015;72(7):542-556. https://doi.org/10.2146/ajhp140280

Accepted 13 July 2017.

Supplementary Table 1 The performance data of the PAWPER XL-MAC system in the derivation and validation stages, shown segment-by-segment.

		Accuracy of derivation model (before calibration)			Accuracy after calibration			Accuracy in validation sample				
	n	PW10	PW20	MPE (ULOA, LLOA)	PW10	PW20	MPE (ULOA, LLOA)	Dataset	n	PW10	PW20	MPE (ULOA, LLOA)
3	-	-	-	-	-	-	-		-	-	-	-
4	7	33.3	83.3	-10.3 (-29.2, 8.5)	100	100	0.0 (-3.4, 3.4)		-	-	-	-
5	17	41.2	76.5	-12.6 (-32.7, 7.4)	88.2	100	-1.9 (-16.3, 12.5)	NHANES data	44	84.1	100	-1.0 (-15.2, 13.2)
6	45	66.7	93.3	-5.6 (-24.2, 12.9)	77.8	100	-1.0 (-15.7, 13.6)		137	81.8	97.1	2.3 (-14.9, 19.6)
7	57	75.0	98.0	-3.3 (-19.5, 12.9)	82.5	98.2	-2.1 (-17.1, 12.9)		219	83.1	98.2	1.0 (-14.8, 16.7)
8	102	78.4	97.1	-3.8 (-18.6, 11.0)	83.3	98.0	-1.5 (-15.9, 12.9)		253	77.5	97.6	4.2 (-18.5, 26.9)
9	105	80.2	99.0	-0.3 (-16.4, 15.7)	85.7	99.0	1.2 (-13.6, 16.1)		295	75.3	98.3	3.6 (-13.0, 20.2)
10	115	81.3	99.1	0.1 (-14.6, 14.9)	86.1	98.3	0.7 (-13.2, 14.5)		319	76.5	97.5	4.4 (-11.6, 20.4)
11	118	76.5	98.3	-1.2 (-17.9, 15.6)	85.6	100	0.5 (-12.8, 13.8)		317	79.2	98.4	4.5 (-9.3, 18.3)
12	139	87.2	99.2	-1.7 (-15.2, 11.8)	92.8	99.3	0.9 (-11.3, 13.2)		418	80.9	97.6	3.9 (-11.7, 19.5)
13	138	75.4	98.5	-6.2 (-18.3, 5.9)	94.2	100	0.1 (-10.7, 10.8)		389	81.0	98.2	3.9 (-10.5, 18.3)
14	118	73.2	98.2	-5.8 (-18.9, 7.3)	92.4	100	-1.6 (-12.8, 9.7)		339	93.2	99.7	0.2 (-11.5, 12.0)
15	109	72.7	99.0	-6.5 (-19.3, 6.4)	87.2	100	-1.2 (-14.0, 11.7)		293	89.1	98.6	0.7 (-12.4, 13.9)
16	93	73.9	94.3	-6.8 (-21.5, 7.9)	90.3	100	-1.8 (-14.0, 10.3)		234	90.2	99.6	1.5 (-11.7, 14.7)
17	87	80.0	95.3	-5.8 (-20.5, 8.8)	94.3	100	-1.6 (-12.8, 9.5)		231	95.7	100	-0.4 (-11.3, 10.4)
18	89	75.0	98.8	-6.1 (-18.7, 6.5)	92.1	100	-2.9 (-12.5, 6.8)		222	94.6	99.5	-1.2 (-12.0, 9.5)
19	96	81.3	97.8	-5.1 (-18.5, 8.2)	90.6	100	-0.9 (-12.6, 10.7)		246	91.1	100	1.1 (-11.6, 13.8)
20	85	64.3	89.3	-8.9 (-24.7, 6.8)	89.4	100	-4.1 (-14.3, 6.1)		278	89.6	99.6	-2.8 (-14.2, 8.5)
22	125	66.7	91.1	-6.9 (-26.5, 12.8)	90.4	98.4	-1.8 (-13.8, 10.3)		350	88.3	98.6	-0.9 (-14.3, 12.5)
24	187	75.1	93.4	-3.6 (-21.3, 14.1)	90.9	99.5	-0.5 (-12.6, 11.5)		431	88.2	99.8	-0.2 (-13.2, 12.8)
26	58	79.3	96.6	-4.0 (-20.3, 12.3)	93.1	98.3	-2.6 (-13.7, 8.5)		369	86.2	98.6	-2.4 (-16.3, 11.4)
28	84	68.7	97.6	-6.1 (-22.4, 10.3)	89.3	98.8	-2.5 (-14.8, 9.8)		149	81.2	98.0	-1.7 (-18.1, 14.7)
30	81	71.8	93.6	-5.3 (-22.0, 11.4)	93.8	100	-2.6 (-13.6, 8.4)		447	85.0	97.5	-2.8 (-19.7, 14.0)

32	91	53.8	84.6	-10.1 (-28.6, 8.4)	83.5	97.8	-5.0 (-17.7, 7.8)		326	86.5	99.1	-3.6 (-16.7, 9.4)
34	114	50.0	85.5	-10.7 (-28.1, 6.8)	81.6	99.1	-4.7 (-17.2, 7.8)		345	80.3	95.9	-4.1 (-19.9, 11.8)
36	115	45.0	82.0	-11.0 (-30.8, 8.7)	85.2	98.3	-3.7 (-16.8, 9.4)		306	80.7	95.4	-4.6 (-21.9, 12.7)
38	141	39.7	65.4	-14.4 (-35.4, 6.5)	87.2	98.6	-3.9 (-17.4, 9.5)		411	84.4	96.8	-3.3 (-20.1, 13.5)
40	161	28.0	60.9	-17.6 (-37.3, 2.2)	78.9	96.9	-5.2 (-19.5, 9.1)		494	79.8	98.2	-2.6 (-18.1, 13.0)
45	534	25.0	73.9	-15.3 (-30.1, -0.4)	80.9	98.3	-1.7 (-18.0, 14.6)		997	81.0	98.9	-1.1 (-17.5, 15.4)
50	546	23.7	75.4	-15.5 (-30.2, -0.9)	80.2	97.4	-2.9 (-19.0, 13.2)		1725	78.3	97.7	-2.4 (-19.2, 14.4)
55	484	39.2	82.5	-13.1 (-31.7, 5.5)	80.6	98.6	-1.7 (-20.4, 17.0)		1039	78.8	97.7	-1.2 (-19.0, 16.7)
60	211	54.3	91.0	-9.7 (-25.7, 6.3)	79.1	98.6	-3.6 (-18.7, 11.6)		678	73.3	97.1	-3.2 (-21.1, 14.8)
65	111	59.6	90.8	-7.9 (-25.0, 9.2)	81.1	97.3	-2.8 (-18.7, 13.1)		378	78.0	95.8	-2.5 (-20.6, 15.6)
70	101	53.9	88.8	-8.9 (-26.5, 8.7)	82.2	98.0	-4.4 (-18.4, 9.6)		454	76.4	97.4	-3.9 (-19.9, 12.1)

Supplementary Table 2 The results of the bootstrap analysis performed on the NHANES data. A bootstrap analysis was conducted because of the large size of the validation dataset. Bootstrapping is a statistical process that relies on random sampling from the original dataset with evaluation of measures of accuracy (bias, variance or other measures) on each subsample. Bootstrapping is usually applied in small samples in an attempt to estimate the true variability of the population from which the sample was drawn. In this study, it was useful to evaluate the impact of the variance on the outcome measures in smaller sample sizes. This is because the effect of many children with normal habitus could mask the effects of children with extremes of habitus. This method therefore allows us to estimate the accuracy of the model in a random selection of children. It would thus be more comparable to other weight estimation studies with similar smaller sample sizes. A sample size of 300 with 100 repetitions was used. The bootstrap best represents the inner limits of agreements and the bootstrap worst represents the outer limits of agreement.

Accuracy in validation sample			
	PW10	**PW20**	**MPE (ULOA, LLOA)**
Bootstrap Mean	82.0	98.1	-1.0 (-14.0, 12.3)
Bootstrap Worst	78.4	96.0	-2.0 (-16.4, 14.5)
Bootstrap Best	85.5	99.8	0.0 (-11.7, 10.1)

9.3 Epilogue

9.3.1 Major findings

This was the original derivation and validation study for the PAWPER XL-MAC system, which provided a model in which mid-arm circumference could be used as part of the weight estimation model without the need to perform any calculations.

The PAWPER XL-MAC model was very accurate in the samples of children from high-income and low- and middle-income populations. It maintained a high level of accuracy and achieved the acceptable outcome criteria in every subgroup except that of severely obese children (BMI-for-age centile >99%).

The accuracy of the PAWPER XL-MAC system in this study was similar to, or better than, historical Mercy method outcomes (PW10 63.9 to 80.1%) [6]. The PAWPER XL-MAC system also proved to be far better than the published accuracy of models in which mid-arm circumference was used as a single variable (PW10 15.5 to 44.2%) [6].

Finally, when evaluating the PAWPER XL-MAC model in the two South African test samples, it achieved a similar accuracy to the outcomes of the regular PAWPER tape used in those studies [7, 8]. The difference in bias, especially in the sample with many underweight children, suggested that the PAWPER XL-MAC system may need to be evaluated more extensively in populations from low- and middle-income countries, which have a greater prevalence of underweight children in their populations.

9.3.2 This study in the context of the thesis

The main focus in this thesis has been the enhancement of the PAWPER system to improve its weight estimation accuracy. There are two elements that influence the ability of a weight estimation system to be able to deliver an accurate weight estimation in the context of paediatric emergency medicine. The first element is the system's innate ability to produce an accurate weight estimation. This initial study confirmed that the PAWPER XL-MAC system was accurate, although additional validation and prospective studies will still be required for confirmation. The

increased objectivity and potential for standardisation is another important advantage of this method, addressing some of the limitations we previously identified. The second element that determines whether a weight estimation system is appropriate for use in emergencies is its simplicity and reliability. While the PAWPER XL-MAC system might ensure that the stress of uncertainty surrounding the assessment of habitus is removed, it might not be suitable for the first weight estimation in emergencies. The time taken to measure mid-arm circumference, especially with adverse patient conditions (such as a child having a seizure), would almost certainly be longer than using a reference image-assisted method. This is an assumption that needs to be tested under research conditions but must be considered. It is possible that an initial, rapid weight estimation could be obtained using the PAWPER XL tape and followed later with the PAWPER XL-MAC system once the early medical management had been accomplished. This will also need to be evaluated in future research. An algorithm for emergency weight estimation could be useful to provide guidance for emergency weight estimation and could also be used to provide structure for simulation teaching and training.

The PAWPER XL-MAC model was found to be accurate at estimating TBW in this study but the mechanism for estimating IBW remains the same as for the PAWPER XL tape and was preserved in this new model [9]. The methodology would still be appropriate to allow for drug dosing in obese children according to best-evidence guidelines.

An additional consideration was whether the tape, which is 40 mm wide, could be used to measure MAC accurately in all children. This will need to be formally tested in further prospective studies. These studies should include an assessment of accuracy, ease-of-use and inter-user reliability in children of different ages.

9.3.3 Link to the next article

The studies in the preceding chapters all showed the PAWPER systems to be accurate and to out-perform many other weight estimation methods. However, these studies were either conducted in databases, in virtual studies, or in cooperative children who did not require emergency treatment. The resilience of a weight

estimation method is crucial to its ability to remain accurate during emergency use. This means that it must remain easy-to-use and remain accurate in adverse physical and cognitive circumstances.

There have been very few studies on weight estimation during simulated emergencies and none during real emergencies [6, 10]. There have also only been a few studies on the effect of human factor errors on the accuracy of weight estimation methods [11, 12]. There have been no studies on the effects of adverse patient cooperation and adverse patient position on the accuracy of weight estimation methods.

The final study in the thesis was one in which the resilience (accuracy and usability) of several weight estimation systems was tested under very adverse circumstances. It also evaluated a mobile phone app that estimated weight based on the PAWPER XL model.

9.3.4 References

[1] Cattermole GN, Leung PY, Mak PS, Graham CA, Rainer TH. Mid-arm circumference can be used to estimate children's weights. Resuscitation. 2010;81(9):1105-10.

[2] Cattermole GN, Graham CA, Rainer TH. Mid-arm circumference can be used to estimate weight of adult and adolescent patients. Emergency Medicine Journal. 2017;34(4):231-236.

[3] Abdel-Rahman SM, Ridge AL. An improved pediatric weight estimation strategy. Open Medical Devices Journal. 2012;4:87-97.

[4] Tanner D, Negaard A, Huang R, Evans N, Hennes H. A prospective evaluation of the accuracy of weight estimation using the Broselow Tape in overweight and obese pediatric patients in the Emergency Department. Pediatric Emergency Care. 2016;33(10):675-678.

[5] WHO Multicentre Growth Reference Study Group. Reliability of anthropometric measurements in the WHO Multicentre Growth Reference Study. Acta Paediatrica. 2006;450(Suppl):38-46.

[6] Wells M, Goldstein L, Bentley A. The accuracy of emergency weight estimation systems in children - a systematic review and meta-analysis. International Journal of Emergency Medicine. 2017;10(29):1-43. https://doi.org/10.1186/s12245-017-0156-5.

[7] Wells M, Goldstein L, Bentley A. A validation study of the PAWPER XL tape: accurate estimation of both total and ideal body weight in children up to 16 years of age. Trauma and Emergency Care. 2017;2(4):1-8.

[8] Georgoulas V, Wells M. The PAWPER tape and the Mercy Method outperform other methods of weight estimation in children in South Africa. South African Medical Journal. 2016;106(9):933-939.

[9] Wells M, Goldstein L. Optimising emergency weight estimation in underweight and obese children: the accuracy of estimations of total body weight and ideal body weight. European Journal of Emergency Medicine. 2018; In Press.

[10] Wells M, Goldstein L, Bentley A. A systematic review and meta-analysis of the accuracy of weight estimation systems used in paediatric emergency care in developing countries. African Journal of Emergency Medicine. 2017;7:S36-S54.

[11] Abdel-Rahman SM, Jacobsen R, Watts JL, Doyle SL, O'Malley DM, Hefner TD, et al. Comparative performance of pediatric weight estimation techniques: a human factor errors analysis. Pediatric Emergency Care. 2017;33(8):548-552.

[12] Marlow RD, Wood DLB, Lyttle MD. Comparing the usability of paediatric weight estimation methods: a simulation study. Archives of Disease in Childhood. 2018.

Chapter 10

The accuracy of paediatric weight estimation during simulated emergencies.

10.1 Prologue

10.1.1 Background

Weight estimation has never been studied during the provision of real-life emergency care, and seldom during simulated emergencies. This is significant because a weight estimation system that is very accurate when tested in a database or on cooperative participants might not be accurate – or usable – in the real-life, frenetic activity of a paediatric resuscitation.

This point has been proven with age-based formulas. Although we now know that these formulas are not accurate even under ideal circumstances, they are also surprisingly vulnerable to human factor errors (errors resulting from the users) [1]. There is strong evidence that these formulas cannot be recalled from memory accurately during emergencies which, if they are not written somewhere immediately accessible, makes them unusable [2-4]. There is also strong evidence that even the simple arithmetic required for these formulas cannot be performed accurately during emergencies [5-7]. This finding has implications to other methods that require any form of arithmetic, such as the Mercy method. Every weight estimation method has potential shortcomings that need to be evaluated during simulated and, preferably, real emergency care. Some sources of error may only be detected when the systems are used under the stressful conditions.

Human factor errors are important, significant and increase during the stress of emergency situations [8, 9]. Even the Broselow tape, which is probably the simplest of all the weight estimation methods, may be used incorrectly in emergencies [10, 11]. No method can be considered to be safe unless it has been researched in realistic conditions and its users trained and well-practised.

For this reason, it was essential to evaluate the PAWPER XL tape under realistic conditions and to identify any problems with its usage that might result in errors when used in real paediatric emergency care.

The development of a mobile phone app which made use of the PAWPER XL methodology for weight estimation and drug dosing information from previously published data was a logical step, given the evidence on cognitive failure during emergency care [5]. An app would also be useful in providing drug dosing information, especially since it would be able to automatically perform all calculations, including the correct and appropriate use of TBW and IBW as dose scalars on obese children. This app was developed on an Android platform during the course of the research. It made use of the PAWPER XL methodology to generate an estimate of weight from an input of length and habitus score (Figure 10-1). It was released free of charge on the PlayStore as the “Emergency Drug Dosing in Children” or “EDD4C” app. This version of the mobile phone app was tested in the study.

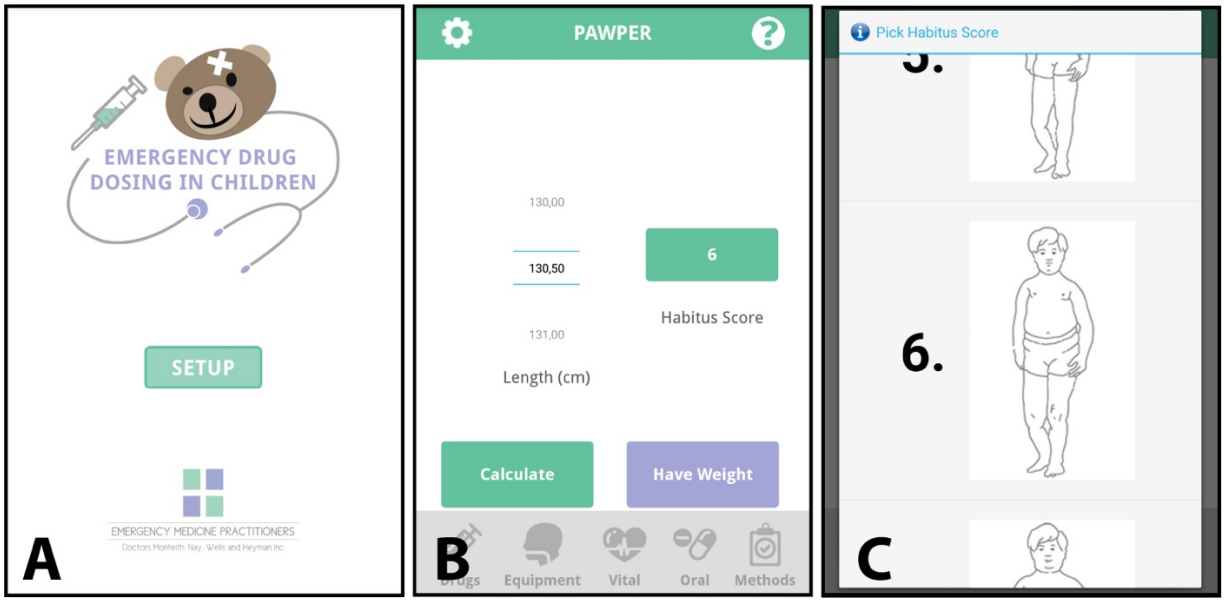


Figure 10-1 Screenshots from the EDD4C app.

10.1.2 Purpose of the study

There have been very few simulation-based studies of weight estimation and none previously with real children in realistic scenarios. The main purposes of this study were to:

- Establish the quantitative accuracy of four weight estimation systems (the Broselow tape, the Mercy method, the PAWPER XL tape and the EDD4C app) in realistic simulated paediatric emergencies, with children in different clinically-appropriate positions and with varying degrees of cooperation.
- Establish the qualitative usability of these four weight estimation systems. In this context, the usability of the weight estimation methods was considered to be characterised by the degree of retention of weight estimation accuracy under adverse patient conditions as well as the least vulnerability to human factor errors (as shown by the difference between the best and worst individual user performances).
- Establish the subjective user-satisfaction with each of the weight estimation systems.

This study was intended to provide practical information on the general applicability of each weight estimation system in a variety of clinical scenarios and, based on the outcomes, whether these systems should be limited to how and when they are used (e.g. only in supine children, not for use in emergencies). Furthermore, user satisfaction would determine the likelihood of whether a system could potentially be adopted in clinical practice and, therefore, insight into the acceptance of the PAWPER and EDD4C systems would assist in further developments and modifications.

This final study concentrated on the questions of the usability of different weight estimation systems contained in Objective 6 (to evaluate the weight estimation accuracy of the PAWPER XL tape system, the EDD4C mobile phone application, the Mercy method and the Broselow tape during simulated paediatric emergencies).

10.2 Published paper

Wells M, Goldstein L, Bentley A. The accuracy of paediatric weight estimation during simulated emergencies: the effects of patient position, patient cooperation and human errors. African Journal of Emergency Medicine. 2018;8(2):43-50.

The PDF of the published article is included. The online-only supplementary material is appended after the PDF. The tables or figures have been retained in the formatting in which they are archived in the journal website.

African Journal of Emergency Medicine 8 (2018) 43–50

Contents lists available at ScienceDirect

African Journal of Emergency Medicine

journal homepage: www.elsevier.com/locate/afjem

ORIGINAL ARTICLE

The accuracy of paediatric weight estimation during simulated emergencies: The effects of patient position, patient cooperation, and human errors

Mike Wells[a,*], Lara Nicole Goldstein[b], Alison Bentley[b]

[a] Division of Emergency Medicine, Faculty of Health Sciences, University of the Witwatersrand, 7 York Road, Parktown, Johannesburg, South Africa
[b] Division of Emergency Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, Gauteng, South Africa

ARTICLE INFO

Keywords:
Body weight
Resuscitation
Paediatric weight estimation
Broselow tape
PAWPER tape
Mercy method

ABSTRACT

Introduction: The effect of patient position and patient cooperation on the accuracy of emergency weight estimation systems has not been evaluated previously. The objective of this study was to evaluate weight estimation accuracy of the Broselow tape, the PAWPER XL tape, the Mercy method, and a custom-designed mobile phone App in a variety of realistic simulated paediatric emergencies.

Methods: This was a prospective study in which 32 emergency medicine volunteers participated in eight simulations of common paediatric emergency conditions, using children models. The participants used each of the four methods to estimate the children's weight. The accuracy of and time taken for the weight estimations were evaluated for each method. A regression analysis determined the effects of patient position and cooperation on weight estimation accuracy. Evaluation of subgroups of best-performers and worst-performers among the participants provided information on the effects of human user-error on weight estimation accuracy.

Results: The Broselow tape, Mercy method, App and the PAWPER XL tape achieved percentages of weight estimation within 10% of actual weight in 47.7, 57.3, 68.1, and 73.0% of estimations, respectively. Patient position and cooperation strongly impacted the accuracy of the Broselow tape, had a minimal effect on the Mercy method and the App, and had no effect on the PAWPER XL tape. The best performing participants achieved very high accuracy with all methods except the Broselow tape.

Discussion: The Mercy method, the App, and the PAWPER XL tape achieved exceptionally high accuracy even in uncooperative and sub-optimally positioned children when used by the best-performing participants. Human error, from inexperience and inadequate training, had the most significant impact on accuracy. The Mercy method was the most subject to human error, and the PAWPER XL tape, the least. Adequate training in using weight estimation systems is essential for paediatric patient safety.

African relevance

- Children in Africa and other low- and middle-income countries are vulnerable to inaccurate weight estimation in emergencies.
- The effect of paediatric patient position on emergency weight estimation is unknown.
- The effect of paediatric patient cooperation on emergency weight estimation is unknown.
- Accurate weight estimation in children in emergencies is feasible with economical equipment options.

Introduction

During the management of paediatric emergencies, errors in drug dosing arising from inaccurate estimations of weight can potentially lead to poor outcomes [1–3]. It is therefore imperative that weight-estimation errors be minimised, so that optimum treatment can be delivered during emergency care. Weight estimation systems that have been proven to be accurate should be used and healthcare providers should be well-trained in their use [4,5]. An accurate weight estimation system, however, may not necessarily lead to accurate weight determination as there are other potential sources of errors that must be considered [6–8]. Health care providers must account for the complexity of the weight estimation system, the experience of the users in weight estimation, and patient factors. These include: inherent limitations of the weight estimation system itself (e.g., age-based weight estimation is not accurate) [9], incorrect use of the device or system [10,11] (even simple systems are susceptible to error when used by

Peer review under responsibility of African Federation for Emergency Medicine.
* Corresponding author.
E-mail address: mike.wells@emergencymedicine.co.za (M. Wells).

https://doi.org/10.1016/j.afjem.2017.12.003
Received 6 June 2017; Received in revised form 7 October 2017; Accepted 13 December 2017
Available online 19 January 2018
2211-419X/ 2018 African Federation for Emergency Medicine. Publishing services provided by Elsevier. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

under-trained individuals [12–14]), or there may be difficulty in using the device or system because of suboptimal patient positioning or uncooperative patients (e.g., the use of a length-based tape in a sitting patient or in a combative, hypoxic child).

Two-dimensional weight estimation systems, which make use of length and habitus to estimate a weight, such as the Mercy method and the PAWPER XL tape, are the most accurate of all the existing weight estimation methodologies [2,15]. However, they have never been evaluated under real or simulated resuscitation conditions. These systems are slightly more complex than simple length-based methods and so, may be more vulnerable to human errors during the stress of emergency care. A weight estimation system that is accurate, but unusable during emergencies, would be of little value.

Advances in technology can also be applied in this setting to reduce errors and cognitive load in emergencies. A mobile-phone App that could rapidly provide accurate point-of-care estimations of weight could be extremely useful. The *Emergency Drug Dosing 4 Children* App was designed to generate estimations of weight, but it has not previously been formally evaluated for patient and human factors [16].

The aim of this study was primarily to evaluate the accuracy of the four weight estimation systems, the Broselow tape, the Mercy method, the PAWPER XL tape, and a weight estimation App, under realistic simulated resuscitation scenarios. We also aimed to establish whether patient factors, including patient positioning and patient cooperation, and participant human error factors, including individual variations in competency, had a substantial impact on the accuracy of the weight estimation.

Methods

This study was a prospective, simulation study conducted in the Emergency Centre of a tertiary, academic hospital in Johannesburg, South Africa. Emergency medicine registrars, emergency medicine consultants, and senior advanced life support paramedics invited to participate in the study. All participants had at least five years' postgraduate experience. Permission to conduct the study was obtained from the Human Research Ethics Committee of the University of the Witwatersrand. All participants signed an informed consent form. An estimated sample size of 30 participants was required to detect a 10% difference in accuracy between the methods based on the McNemar test, powered to 80% at a 0.05 significance level and assuming a baseline accuracy of 70%.

Eight simulation stations were prepared, each recreating a commonplace emergency scenario, with a child volunteer who simulated the medical condition, patient position, and degree of cooperation specified for the scenario (Table 1). The scenarios were chosen to recreate a spectrum of common realistic medical circumstances in which children might not be supine and not cooperative. Some children were fully cooperative, some simulated actively uncooperative patients, such as with a seizure, and some were passively uncooperative, such as during a cardiac arrest. The children were positioned in clinically realistic positions, for example, a supine position for a child in cardiac arrest and a sitting position for a child in respiratory distress.

All participants attended a formal one-hour training session with the weight estimation methods and had an opportunity to practice to their satisfaction before starting the simulation. At each station, every participant used all four techniques to estimate weight. The sequence of methods used was different in each station, based on a randomised, predetermined order. The participants were instructed to obtain each weight estimate independently, not taking the results of the other methods into account. They were informed that the estimations might not be the same between the different methods. They were blinded to the measured weight of the children and the weight estimates of other participants. The time taken to estimate the weights was recorded by a time-keeper.

Upon completion of the six minutes allocated for each station, the participants rotated to the next station until all eight stations were completed. The participants then completed a questionnaire on their experiences with the different systems. Finally, the investigators, experts in the use of the weight-estimation methods, used each system in each child model, now fully cooperative, to provide control data for weight estimation accuracy under ideal circumstances.

Four weight estimation methods were tested in this study (Fig. 1): the PAWPER XL tape [2,17], the Broselow tape 2011 edition A, the Mercy method [15] and the *Emergency Drug Dosing 4 Children* mobile phone App.

Table 1
Details of simulation scenarios, simulated patients and the demographic characteristics of the participants.

Station number	Characteristics of simulated patient	Clinical scenario including position of child and cooperation	Order of weight estimation methods
1	Age: 9 years Weight: 22.5 kg Length: 130 cm HS: 1 (very underweight)	***Cardiac arrest*** Supine on bed Floppy, no cooperation or resistance	PTXL, APP, MM, BT
2	Age: 12 years Weight: 39.6 kg Length:151 cm HS: 4 (overweight)	***Status epilepticus*** Lateral position on bed Arms and legs jerking, no cooperation	APP, MM, BT, PTXL
3	Age: 11 years Weight: 34.2 kg Length: 145 cm HS: 3 (average)	***Respiratory distress – severe asthma*** Sitting on bed. Will not lie down Anxious, semi-cooperative	BT, PTXL, APP, MM
4	Age: 13 years Weight: 55.8 kg Length: 156 cm HS: 5 (obese)	***Major trauma with abdominal injuries*** Immobilised supine on stretcher Cooperative	MM, PTXL, BT, APP
5	Age: 16 years Weight: 61.2 kg Length: 174 cm HS: 3 (average)	***Major trauma with head injury GCS 10/15*** Supine on spine board Uncooperative, irritable	MM, BT, PTXL, APP
6	Age: 1 year Weight: 11.7 kg Length: 80 cm HS: 4 (overweight)	Severe gastroenteritis with hyperkalaemia Sitting on mom's lap Not cooperative	APP, MM, BT, PTXL
7	Age: 8 years Weight: 31.2 kg Length: 132 cm HS: 4 (overweight)	***Unstable supraventricular tachycardia*** Semi-recumbent on bed Fully cooperative	PTXL, MM, APP, BT
8	Age: 7 years Weight: 26.4 kg Length: 131 cm HS: 2 (underweight)	***Severe pneumonia with hypoglycaemia*** Sitting down. Floppy. Semi-cooperative	BT, APP, PTXL, MM

Characteristics of participants

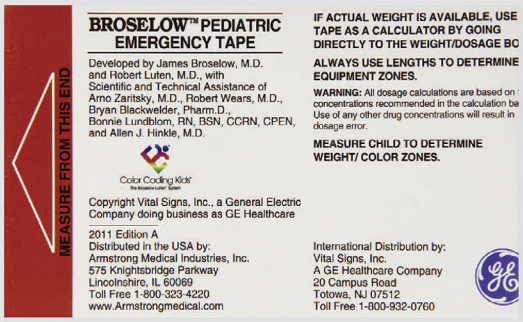
Qualification	N	Sex – male n (%)	Experience Years median (LQ, UQ)	Confidence with children score median (LQ, UQ)
All	32	21 (65.6)	5.5 (3.0, 11.5)	6 (4, 7)
Emergency medicine registrars	21	14 (66.7)	5.0 (4.0, 6.8)	6 (4, 7)
Emergency medicine consultants	5	1 (20.0)	9.0 (7.5, 9.8)	7 (7, 7)
Senior advanced life support paramedics	6	6 (100)	12.5 (11.3, 15.3)	7 (7, 8)

HS, habitus score; PTXL, PAWPER XL tape; MM, Mercy method; BT, Broselow tape; LQ, lower quartile; UQ, upper quartile.

The Broselow tape

The Broselow tape is used to measure a child from head to heel – the weight is read directly off the tape at that point.

Limitations: the tape is only 145cm in length, so children taller than the tape must receive either an adult weight estimate or the highest tape weight (36kg).



The Mercy method

Step 1: humerus length is measured in the seated child from the inferior aspect of the acromion to the tip of the olecranon. The Mercy table is used to look up the partial weight that corresponds to this length. The midpoint of the humerus is also marked. Step 2: mid-arm circumference is measured at the midpoint of the arm and a partial weight obtained from the table. Step 3: the two partial weights are added to obtain the Mercy method prediction of total body weight.

Limitations: two careful measurements are used and basic arithmetic is required to obtain a weight estimate.

Humeral Length (cm)	Partial Weight A (kg)	Mid-upper arm Circumference (cm)	Partial Weight B (kg)
9	0.5	10	2.8
10	0.7	11	3.8
11	0.9	12	4.6
12	1.5	13	4.9
13	2.0	14	5.3
14	2.8	15	5.9
15	3.4	16	6.5

The PAWPER XL tape

The PAWPER XL tape is the second generation of the PAWPER tapes - it is longer than the original tape (180cm vs 153cm) and has additional habitus scores for obese children. Step 1: The tape is used to measure a child from head to heel – the colour-zone and weight-zone is read directly off the tape at that point. Step 2: A visual assessment, or an image-assisted assessment of the child's habitus is made and a habitus score assigned – HS1 is very thin, HS2 is slightly underweight, HS3 is average, HS4 is slightly overweight, HS5 is obese, HS6 and HS7 represent severe obesity. The adjusted weight can be read off the tape for the appropriate habitus score.

Limitations: estimation of habitus is still subjective unless the figural image-assisted method is used.



The EDD4C App

The Emergency Drug Dosing 4 Children (EDD4C) app can generate estimations of weight emulating the PAWPER XL tape methodology. Step 1: The child's length is measured using a regular measuring tape. This length is entered into the App using the scrollbar. Step 2: The user visually assesses the child's habitus and then uses an image on the app to select the habitus score. The App will automatically calculate the adjusted weight.

Limitations: this version of the app required an active internet connection.

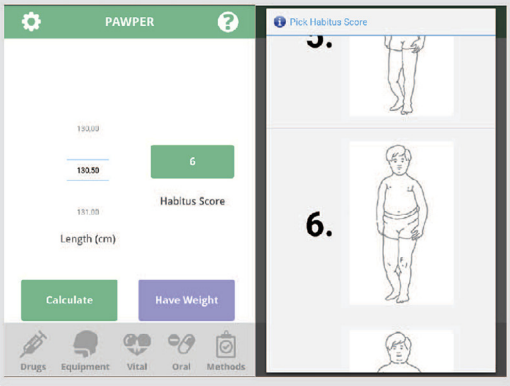


Fig. 1. Details of the weight-estimation systems used in this study. Descriptions of how the methods are used are provided. Abbreviations: EDD4C app (emergency drug dosing for children mobile application).

The performance of weight estimation of each system was evaluated for accuracy, bias and precision. Accuracy was represented by the percentage of estimations falling within 10% and 20% of actual weight (PW10 and PW20). The PW10 is a method of quantifying the overall accuracy of a weight estimation system and takes both bias and precision into account. An acceptable accuracy for a weight estimation system has been suggested to be a PW10 > 70% [18]. The PW20 provides a measure of overall accuracy, but also provides an indication of

the critical errors produced by a weight estimation system: those with a > 20% error. An acceptable critical error rate for a weight estimation system has been suggested to be 5% (PW20 > 95%) [19]. The bias was represented by the mean percentage error of estimations and the precision by the 95% limits of agreement of the percentage error; these were the Bland & Altman limits of agreement i.e., MPE ± 1.96 × standard deviation of percentage error. We determined the weight estimation accuracy achieved by each individual participant and calculated the accuracy data for the subgroups of best-performing and worst-performing participants. We then compared the systems based on overall accuracy (PW10 data) using the McNemar test for paired non-parametric data; we calculated odds ratios for the significant outcomes. We used the Mann-Whitney test to make comparisons between the times taken to complete the estimations.

In order to identify the effects of patient characteristics on the accuracy of weight estimation, we performed a multiple logistic regression analysis for each of the systems, with weight estimation accuracy, defined as estimates within 10% of actual weight, as the dependent variable and patient length, habitus score, position and cooperation as the independent variables. We quantified the effect sizes of statistically significant findings using crude and adjusted odds ratios. We used the R^2 to measure how much variability was explained by each model.

The primary outcome measure was weight estimation accuracy (based on PW10 data). The secondary outcome measures were the time taken to complete weight estimation and feedback from the participants on the use of the systems.

A $p < 0.05$ was regarded significant for all statistical tests, including the predictors in the multivariate models. Statistical analysis was performed using Stata (StataCorp, 2015, Stata Statistical Software: Release 14, College Station, TX: StataCorp LP).

Results

Thirty-two participants were enrolled in the study with a variety of demographics (Table 1).

This study determined the accuracy of weight estimation for each method, the experts' control data, and the times taken for weight estimation (Table 2). The PAWPER XL tape and the App had the highest accuracy, least bias, and best precision. The PAWPER XL tape and the App also had the lowest critical error rate, 4.8% and 5.0% respectively, and the Broselow tape the highest at 34.4%. This pattern was repeated with the experts' weight estimation control data (Fig. 2).

The data for the weight estimation accuracy by the participants and the expert "controls" were calculated with the groups of best and worst performers in each system (Table 2). A negative figure in the mean percentage error and percentage error limits of agreement indicates an underestimation of weight (Table 2). Statistical tests used for the comparisons were the McNemar test for paired outcome comparisons and the Mann-Whitney test for paired time-data. The superscripts designate significant differences in the comparisons between the weight estimation systems; for example, in the Broselow tape column, the PW10 data "$47.7^{ed,mm,pt}$" indicates a significant difference in the comparisons between the PW10 of the Broselow tape^{bt}, the EDD4C app^{ed}, the Mercy method^{mm} and the PAWPER XL tape^{pt} (Table 2). Statistical testing was not performed on the mean percentage error (MPE) data as statistically significant differences in bias provided limited information.

There was substantial variation between individual participants in the accuracy of weight estimation achieved. The highest-performers amongst the participants achieved exceptionally good accuracy with the Mercy method, the App and the PAWPER XL tape, achieving PW10s of 78.6%, 80.4% and 83.6% respectively. The worst-performers were least accurate with the Mercy method when compared with the App and the PAWPER XL tape, achieving PW10s 36.2%, 55.6% and 62.5% respectively.

The Broselow tape and PAWPER XL tape had significantly faster estimation times than the App and Mercy method, with about 80% of estimations completed within 30 s. Even with the slower methods, weight estimation was completed in less than one minute in more than 75% of scenarios.

Multiple regression analysis were also performed (Table 3). The accuracy of the PAWPER XL tape was unaffected by the children's length, habitus, position or degree of cooperation, the Mercy method and the App were slightly affected, but the Broselow tape was significantly impacted by these factors. For the Broselow tape, habitus deviations from the average were associated with the largest impact on accuracy. Position and cooperation had no effect on the time to obtain weight estimates for any system, but taller and overweight children had slightly slower weight estimation times with the App.

Discussion

This is the first study to have evaluated the Mercy method and the PAWPER XL tape in a simulated resuscitation setting. It is also the first to have evaluated a weight-estimation App and the first study to have evaluated the effects of suboptimal patient positions and behaviour on weight estimation accuracy.

When considering the accuracy of the systems, the Broselow tape was not ideal according to the results of this study. Its performance was similar to that reported in many previous articles with PW10s between 45% and 60%, however its accuracy was better than its performance in very underweight populations with PW10s between 10% and 30% [20–23]. Both its poor accuracy and high critical error rate continue to raise doubts about its safety when used to guide drug dose calculations in many populations. The Mercy method has previously been shown to be very accurate in populations from both high-income and low- and middle-income countries across the world with PW10 70–80% [2,24–28]. In this study, its accuracy was substantially lower with a PW10 of 57%, but it was comparable to the accuracy reported when the method was used by non-experts with PW10 46–53% [11,24]. The PAWPER XL tape performed very well in this study with a PW10 73.0%, similar to the results from the only other previous study on the PAWPER XL tape [17]. The App achieved a moderate degree of weight estimation accuracy, but was slightly less accurate than the PAWPER XL tape-system on which it is based. This data was encouraging, however, and supports the value of further research into weight estimating point-of-care mobile phone apps, in addition to the tape-based PAWPER XL system. The use of an App that could provide an accurate estimate of weight as well as comprehensive information on drug dosing would be enormously useful.

The accuracy data from the experts illustrated the excellent levels of accuracy the two-dimensional weight estimation systems could achieve in an ideal scenario. The difference in accuracy between the experts and the participants was a strong indication that patient-induced errors, such as from patient positioning or lack of cooperation, or user-induced errors impact weight estimation accuracy, more than an inaccuracy of the methodology itself [11,29].

The effects of length and habitus were important for the Broselow tape: the tape was significantly and substantially less accurate in underweight and overweight children, a major disadvantage common to all length-based weight estimation systems [9,30]. It is also the reason that the Broselow tape has been repeatedly shown to be inaccurate in populations with a high prevalence of underweight or obese children, or both [31]. Length plus a measure of habitus is required for an accurate weight estimation [32]. Taller children also had a reduced accuracy of weight estimation with the Broselow tape. Since the length limitations of the tape are quite restrictive at 145 cm, children as young as nine or ten years of age may not receive accurate weight estimations [33]. While children "too tall for the tape" are often excluded during research studies, this is not possible in the real world. In this study, two children were too tall for the tape and had substantial underestimation of their weight, despite the participants following the manufacturers'

Table 2
Performance of the four weight estimation systems tested in the simulations.

Weight-estimation accuracy data of participants				
Participants	Broselow tape	EDD4C app	Mercy method	PAWPER XL tape
n	256	235	253	256
Mean percentage error (%)	−10.0	2.8	−6.8	−1.8
Limits of agreement (%)	−50.8, 30.8	−25.4, 31.1	−35.2, 21.5	−23.6, 19.9
PW10 (%)	47.7[ed,mm,pt]	68.1[bt]	57.3[bt,pt]	73.0[bt,mm]
PW20 (%)	65.6[ed,mm,pt]	95.0[bt]	85.8[bt,pt]	95.2[bt,mm]
Weight-estimation accuracy data of experts				
Experts	Broselow tape	EDD4C app	Mercy method	PAWPER XL tape
n	24	24	24	24
Mean percentage error (%)	−10.0	3.3	−6.0	−1.5
Limits of agreement (%)	−50.8, 30.8	−7.0, 13.5	−18.4, 6.5	−15.0, 12.0
PW10	50.1[ed,mm,pt]	87.5[bt]	87.5[bt]	87.5[bt]
PW20	62.6[ed,mm,pt]	100[bt]	95.8[bt]	100[bt]
Stratified weight-estimation accuracy of participants – accuracy (PW10) data				
Participants	Broselow tape	EDD4C app	Mercy method	PAWPER XL tape
n	32	32	32	32
Top 10 (best in group) (%)	56.4[ed,mm,pt]	84.0[bt]	82.3[bt]	87.5[bt]
Bottom 10 (worst in group) (%)	36.3[ed,pt]	51.4[bt,mm]	24.1[ed,pt]	57.5[bt,mm]
Top half of group (%)	53.9[ed,mm,pt]	80.4[bt]	78.6[bt]	83.6[bt]
Bottom half of group (%)	41.4[ed,pt]	55.6[bt,mm]	36.2[ed,pt]	62.5[bt,mm]
Time to complete weight estimation				
Participants	Broselow tape	EDD4C app	Mercy method	PAWPER XL tape
n	256	235	253	256
Time (s)–median (LQ, UQ)	20 (14, 26)[ed,mm,pt]	44 (33, 60)[bt,pt]	44 (37, 54)[bt,pt]	23.5 (18, 29)[bt,ed,mm]
Time < 30 s – n (%)	218 (85.2)[ed,mm]	43 (17.0)[bt,mm,pt]	23 (9.0)[bt,ed,pt]	203 (79.3)[ed,mm]
Time > 60 s – n (%)	2 (0.8)[ed,mm]	66 (25.7)[bt,pt]	36 (14.1)[bt,pt]	2 (0.8)[ed,mm]

bt, Broselow tape; ed, EDD4C app; mm, Mercy method; pt, PAWPER XL tape.

guidelines for this scenario to use an adult weight.

Length and habitus had a far smaller impact on the accuracy of the Mercy method and the App, and none at all on the PAWPER XL tape, because the two-dimensional systems take both length and habitus into account in the estimation methodology [2,15]. The observed associations were unlikely to significantly impact on overall accuracy. Previous studies on the original PAWPER tape showed a high accuracy in populations with a low prevalence of obesity with a PW10 > 80%, but a lower accuracy in populations with a high incidence of obesity with a PW10 60–70% [2,24,34,35]. The lack of association with habitus

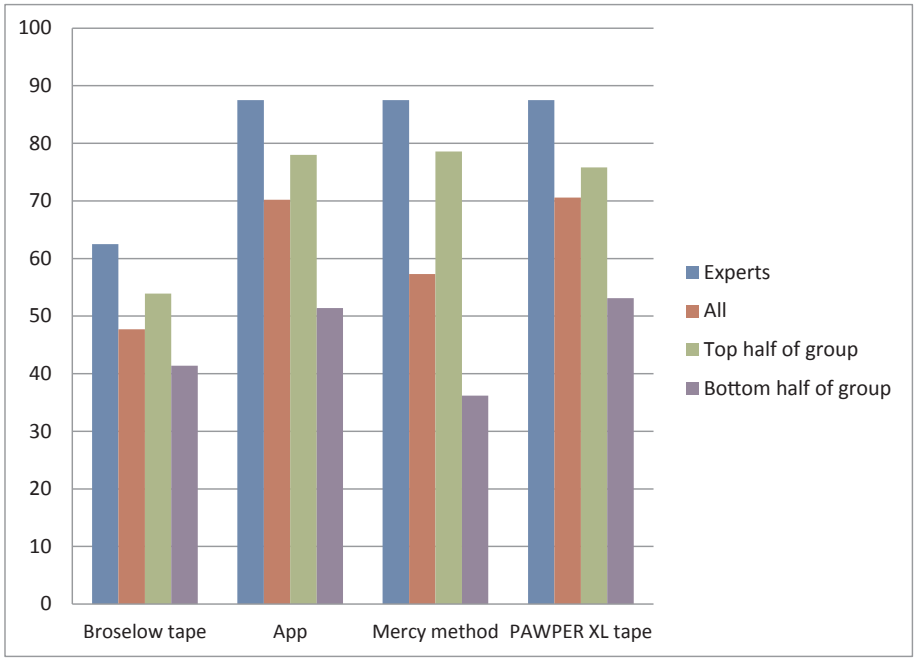


Fig. 2. Accuracy of weight estimation of each of the systems. The histogram shows the weight estimation accuracy for each weight estimation system for the pooled participants, the experts, the best 16 participants and the worst 16 participants (top and bottom half of the group, respectively). Percentages of weight estimates within 10% of actual weight (PW10) are shown.

Table 3
The results of the multiple regression analyses for weight estimation accuracy and time to obtain a weight estimation.

Accuracy of weight estimation							
	Broselow tape		Mercy method		App		PAWPER XL tape
Position	Crude OR	6.5 (3.0, 10)	Crude OR	0.4 (0.4, 0.5)	Crude OR	2.5 (1.4, 4.4)	NS
	Adjusted OR	5.9 (1.9, 26)	Adjusted OR	0.6 (0.5, 0.7)	Adjusted OR	3.8 (1.4, 10)	
	Crude R^2	0.21	Crude R^2	0.35	Crude R^2	0.25	
	Adjusted R^2	0.65	Adjusted R^2	0.40	Adjusted R^2	0.32	
	Interpretation: BT more accurate in supine children		Interpretation: MM more accurate in sitting children		Interpretation: MM more accurate in supine children		
Cooperation	Crude OR	4.9 (2.9, 8.3)	Crude OR	2.7 (2.2, 3.2)	NS		NS
	Adjusted OR	1.6 (1.2, 2.7)	Adjusted OR	3.6 (1.6, 8.2)			
	Crude R^2	0.26	Crude R^2	0.25			
	Adjusted R^2	0.65	Adjusted R^2	0.40			
	Interpretation: BT more accurate in cooperative children		Interpretation: MM more accurate in cooperative children				
Length	Crude OR	5.8 (4.0, 7.6)	Crude OR	1.1 (1.0, 1.3)	Crude OR	1.1 (1.0, 1.8)	NS
	Adjusted OR	4.8 (3.2, 7.4)	Adjusted OR	1.2 (1.0, 1.4)	Adjusted OR	1.3 (1.1, 2.5)	
	Crude R^2	0.19	Crude R^2	0.15	Crude R^2	0.14	
	Adjusted R^2	0.65	Adjusted R^2	0.40	Adjusted R^2	0.32	
	Interpretation: BT less accurate in taller children		Interpretation: MM less accurate in taller children		Interpretation: EDD4C less accurate in taller children		
Habitus	Crude OR	28 (19, 37)	Crude OR	1.1 (1.0, 1.4)	Crude OR	1.4 (1.1, 1.8)	NS
	Adjusted OR	18 (14, 22)	Adjusted OR	1.1 (1.0, 1.3)	Adjusted OR	1.4 (1.0, 1.8)	
	Crude R^2	0.46	Crude R^2	0.10	Crude R^2	0.12	
	Adjusted R^2	0.65	Adjusted R^2	0.40	Adjusted R^2	0.32	
	Interpretation: BT less accurate in underweight, overweight and obese children		Interpretation: MM less accurate in obese children		Interpretation: BT less accurate in underweight children		

Time to obtain weight estimate					
	Broselow tape	Mercy method	App		PAWPER XL tape
Position	NS	NS	NS		NS
Cooperation	NS	NS	NS		NS
Length	NS	NS	Crude OR	1.2 (1.1, 1.3)	NS
			Adjusted OR	1.3 (1.1, 1.5)	
			Crude R^2	0.23	
			Adjusted R^2	0.19	
			Interpretation: EDD4C slower in taller children		
Habitus	NS	NS	Crude OR	(1.0, 1.3)	NS
			Adjusted OR	(1.1, 1.4)	
			Crude R^2	0.17	
			Adjusted R^2	0.19	

OR, odds ratios; PTXL, PAWPER XL tape; EDD4C; *Emergency Drug Dosing 4 Children* mobile application; MM, Mercy method; BT, Broselow tape.

in this study provided preliminary evidence that the PAWPER XL tape may have addressed this problem.

Suboptimal patient positioning and lack of patient cooperation negatively influenced the performance of the Broselow tape considerably, which was far more evident than in the other weight estimation systems. The tentative explanation for this is that the width of the tape at 95 mm may have made it cumbersome and difficult to manipulate in sitting or uncooperative children.

There was a similar, although smaller, association with the Mercy method: it was slightly less accurate in uncooperative children, but it was more accurate in sitting children, while the opposite was true for the App. Since the best performers amongst the participants demonstrated little difference in accuracy to that of the experts using the systems under ideal circumstances, the net adverse effects of position and cooperation were likely to be a decrease in accuracy of only a few percentage points at most.

There was no association between patient position or behaviour and the performance of the PAWPER XL tape. The negligible difference between the best participants' accuracy and the experts' accuracy under ideal circumstances confirmed the resilience of this method in this respect. The smaller width of the tape of 40 mm, when compared with the Broselow tape, may have made it easier to manipulate in uncooperative or sitting patients.

When considering the effects of human factor errors, the results achieved by the experts proved that it was possible to achieve a high level of weight estimation accuracy in these children with the Mercy method, the PAWPER XL tape, and the App. Furthermore, the subgroup of best-performers amongst the participants proved that it was possible to achieve a high degree of accuracy despite the adverse patient factors. Human error factors were therefore far more significant and important: it is likely that training, experience and inherent skill played the major role in determining the final accuracy of weight estimation. This has been shown previously in a human factor errors analysis on the Mercy method, which showed very poor accuracies in 80 minimally-trained users [11]. In this study, the largest disparity between the accuracy of the best-performing and worst-performing participants was seen with the Mercy method, which suggests that this more complex technique was the most vulnerable to human error. This gap, while still present, was not as marked with PAWPER XL tape and the App. Differences in accuracy between experienced and inexperienced users have been shown in other studies with the PAWPER tape, although even brief formal training sessions with users eliminated these errors [2,35,36].

Even the Broselow tape, which appears to be simple to use, has been found to be dependent on training: errors are extremely common unless

users have been formally trained with the tape [10,37].

The very positive aspect of these findings is that appropriate training is likely to enable high levels of accuracy to be consistently reproduced in real-life practice with the PAWPER XL tape and the Mercy method [38]. Proper training in weight estimation for use in emergencies should not be ignored and should probably be included as part of paediatric advanced life support courses as well as routine paediatric simulation training.

The Broselow tape and the PAWPER XL tape were both successful at providing a rapid estimate of weight. While this finding was expected for the Broselow tape, in the case of the PAWPER XL tape, it confirmed that habitus assessment was a quick process and therefore probably added minimal cognitive stress. The Mercy method, however, was significantly slower. The multiple measurements and calculations required were the cause of this delay. The use of the App was relatively slow because it was hindered by a scrollbar data-entry system that was very time-consuming to use. Despite the relative slowness of these two systems, however, they did not take longer than 60 s to complete. Although there is no published benchmark for weight estimation times, these systems were all satisfactorily efficient.

One of the most interesting findings of the study was that many of the participants realised that they required more training and practice in weight estimation than they had expected. Although they had considered themselves to be well-trained, they discovered during the simulations that they did not have enough practice.

The Mercy method was found to be complex and too difficult to use in an emergency but, despite its limitations, the App was well-liked. The PAWPER XL tape was regarded by most participants as the most appropriate of the methods for clinical emergency usage.

There were technical problems with fluctuating Wi-Fi access for the *Emergency Drug Dosing 4 Children App* within the hospital precincts on the day of the study. This slowed its functioning and may have skewed data against the App.

Performing sequential weight estimations could potentially have introduced bias, but this was addressed by rotating the order of estimations. A previous study using a similar protocol, however, did not demonstrate any bias from sequential weight estimations [11].

We did not evaluate the participants' previous experience with weight estimation methods. Since some methods are commonly used in South Africa, such as the Broselow tape, these methods may have been more familiar to the participants.

The sample of children in this study was skewed to an older age-group, so the weight estimation accuracy findings might not be generalizable to a younger population. The findings with regards to patient-induced errors and user-induced errors might not remain true and relevant in all age-groups, however. Most weight estimation systems are more accurate in younger children, which may also have biased the results towards those systems, such as the PAWPER XL tape, which are accurate across a broad range of age-groups.

The Broselow tape performed poorly in this study, as in many others before, and it is difficult to see an ongoing future role for the tape when compared to other, newer weight estimation systems. It was the method most affected by suboptimal patient positioning and behaviour. The accuracy of the Mercy method was only slightly impacted by adverse patient positioning and behaviour, but was the system most affected by human factor errors. The App achieved reasonable accuracy, but was the slowest of the methods evaluated because of design limitations. It was only slightly affected by patient factors, but was impacted by differences in ability between participants. The App was popular with the participants but needs further research. The PAWPER XL tape was completely unaffected by patient factors and was the most resilient and least vulnerable to human factor errors, although these still had a significant impact on its overall performance.

While patient factors frequently cannot be safely changed to improve weight estimation accuracy, human factor errors can be reduced through training. Paediatric weight estimation training should become an integral part of resuscitation simulation teaching.

Conflicts of interest

MW designed and created the PAWPER tape that is reviewed in this manuscript. MW is also an editor of this journal. He was, however, not involved in the peer review, decision or editing process of this paper. There are no further conflicts of interest.

Authors' contributions

MW conceived the original idea. MW and LG designed the scenarios and collected the data. MW carried out the analysis of data. MW drafted the manuscript; LG and AB revised it. MW, LG and AB approved the final version that was submitted.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.afjem.2017.12.003.

References

[1] Cattermole GN, Graham CA, Rainer TH. Paediatric weight estimation. Emerg Med J 2012;29(1):82–3.
[2] Wells M, Coovadia A, Kramer E, Goldstein L. The PAWPER tape: a new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. Resuscitation 2013;84(2):227–32.
[3] Srinivasan V, Nadkarni VM, Helfaer MA, Carey SM, Berg RA. American Heart Association National Registry of Cardiopulmonary Resuscitation I. Childhood obesity and survival after in-hospital pediatric cardiopulmonary resuscitation. Pediatrics 2010;125(3):e481–8.
[4] Luscombe M. "Kid's aren't what they used to be": a study of paediatric patients' weights and their relationship to current weight estimation formulae. Br J Anaesth 2005;95(4):578.
[5] Lerner C, Gaca AM, Frush DP, Hohenhaus S, Ancarana A, Seelinger TA, et al. Enhancing pediatric safety: assessing and improving resident competency in life-threatening events with a computer-based interactive resuscitation tool. Pediatr Radiol 2009;39(7):703–9.
[6] Luten R, Wears RL, Broselow J, Croskerry P, Joseph MM, Frush K. Managing the unique size-related issues of pediatric resuscitation: reducing cognitive load with resuscitation aids. Acad Emerg Med 2002;9(8):840–7.
[7] Porter E, Barcega B, Kim TY. Analysis of medication errors in simulated pediatric resuscitation by residents. West J Emerg Med 2014;15(4):486–90.
[8] Kaufmann J, Laschat M, Wappler F. Medication errors in pediatric emergencies: a systematic analysis. Deutsches Ärzteblatt International 2012;109(38):609–16.
[9] Wells M, Goldstein L, Bentley A. It is time to abandon age-based emergency weight estimation in children! A failed validation of 20 different age-based formulas. Resuscitation 2017;116(7):73–83.
[10] Hohenhaus SM, Frush KS. Pediatric patient safety: common problems in the use of resuscitative aids for simplifying pediatric emergency care. J Emerg Nurs 2004;30(1):49–51.
[11] Abdel-Rahman A, Jacobsen R, Watts J, Doyle S, O'Malley D, Hefner T, et al. Comparative performance of pediatric weight estimation techniques: a human factor errors analysis. Pediatr Emerg Care 2015.
[12] Hohenhaus SM. Assessing competency: the Broselow-Luten resuscitation tape. J Emerg Nurs 2002;28(1):70–2.
[13] Shah AN, Frush K, Luo X, Wears RL. Effect of an intervention standardization system on pediatric dosing and equipment size determination: a crossover trial involving simulated resuscitation events. Arch Pediatr Adolesc Med 2003;157(3):229–36.
[14] Lammers R, Willoughby-Byrwa M, Fales W. Medication errors in prehospital management of simulated pediatric anaphylaxis. Prehosp Emerg Care 2014;18(2):295–304.
[15] Abdel-Rahman SM, Ridge AL. An improved pediatric weight estimation strategy. Open Med Devices J 2012;4:87–97.
[16] Emergency drug dosing 4 children [Internet]. 2014. Available from: https://play.google.com/store/apps/details?id=emergencyapp.emergencyapp.
[17] Wells M, Goldstein L, Bentley A. A validation study of the PAWPER XL tape: accurate estimation of both total and ideal body weight in children up to 16 years of age. Trauma Emerg Care 2017;2(4):1–8.
[18] Wells M, Goldstein L, Bentley A. High-tech adjuncts to emergency weight estimation: Point-of-care ultrasound and point-of-care bioelectrical impedance measurements can increase the accuracy of length-based weight estimation in children. Trauma Emerg Care 2017;2(3):1–8.
[19] Stewart D. Accuracy of the Broselow tape for estimating paediatric weight in two Australian Emergency Departments. University of Sydney; 2009.
[20] Graves L, Chayen G, Peat J, O'Leary F. A comparison of actual to estimated weights

in Australian children attending a tertiary children's hospital, using the original and updated APLS, Luscombe and Owens, Best Guess formulae and the Broselow tape. Resuscitation 2014;85(3):392–6.
[21] Suh D, Kwak Y, Kim D, Jeong J. Estimating weights using the Broselow tape in Korea: comparison of old and new versions of the Broselow tape use. Ann Emerg Med 2013;62(4):S109.
[22] Geduld H, Hodkinson PW, Wallis LA. Validation of weight estimation by age and length based methods in the Western Cape, South Africa population. Emerg Med J 2011;28(10):856–60.
[23] Clark MC, Lewis RJ, Fleischman RJ, Ogunniyi AA, Patel DS, Donaldson RI. Accuracy of the Broselow Tape in South Sudan, "The Hungriest Place on Earth". Acad Emerg Med 2016;23(1):21–8.
[24] Georgoulas V, Wells M. The PAWPER tape and the Mercy Method outperform other methods of weight estimation in children in South Africa. S Afr Med J 2016;106(9):933–9.
[25] Batmanabane G, Kumar Jena P, Dikshit R, Abdel-Rahman SM. Using the Mercy Method for weight estimation in Indian children. Global Pediatr Health 2015:1–7.
[26] Dicko A, Alhousseini ML, Sidibé B, Traoré M, Abdel-Rahman SM. Evaluation of the Mercy weight estimation method in Ouelessebougou, Mali. BMC Public Health 2014;14:270.
[27] Abdel-Rahman SM, Paul IM, James LP, Lewandowski A. Evaluation of the Mercy TAPE: performance against the standard for pediatric weight estimation. Ann Emerg Med 2013;62(4). 332–339.e6.
[28] Abdel-Rahman SM, Ahlers N, Holmes A, Wright K, Harris A, Weigel J, et al. Validation of an improved pediatric weight estimation strategy. J Pediatr Pharmacol Ther 2013;18(2):112–21.
[29] Hunt EA, Hohenhaus SM, Luo X, Frush KS. Simulation of pediatric trauma stabilization in 35 North Carolina emergency departments: identification of targets for performance improvement. Pediatrics 2006;117(3):641–8.
[30] Yamamoto LG, Inaba AS, Young LL, Anderson KM. Improving length-based weight estimates by adding a body habitus (obesity) icon. Am J Emerg Med 2009;27(7):810–5.
[31] Young KD, Korotzer NC. Weight estimation methods in children: a systematic review. Ann Emerg Med 2016;68(4). 441–451.e10.
[32] Losek JD, Garland J, Nelson DB. Body habitus + height = accurate weight estimate. Ann Emerg Med 1989;18(6):709.
[33] Cattermole GN, Leung PY, Graham CA, Rainer TH. Too tall for the tape: the weight of schoolchildren who do not fit the Broselow tape. Emerg Med J 2013. [epub ahead of print].
[34] Chavez H, Peterson R, Lo K, Arel M. Weight estimation in an inner city Pediatric Emergency Department: the effect of obesity. Am J Emerg Med 2015;33(10):1364–7.
[35] Garcia CM, Meltzer JA, Chan KN, Cunningham SJ. A validation study of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape-a new weight estimation tool. J Pediatr 2015;167(1). 173–177.e1.
[36] O'Leary F, John-Denny B, McGarvey K, Hann A, Pegiazoglou I, Peat J. Estimating the weight of ethnically diverse children attending an Australian emergency department: a prospective, blinded, comparison of age-based and length-based tools including Mercy, PAWPER and Broselow. Arch Dis Child 2017;102(1):46–52.
[37] Hohenhaus SM, Frush KS. Pediatric patient safety in resuscitation simulations: lessons learned from studying clinician behavior. Pediatr Emerg Care 2005;21(1):71–5.
[38] Frush K, Hohenhaus S, Luo X, Gerardi M, Wiebe RA. Evaluation of a Web-based education program on reducing medication dosing error: a multicenter, randomized controlled trial. Pediatr Emerg Care 2006;22(1):62–70.

Supplementary Table Representative responses and key themes obtained from the post-participation survey. The number of multiple similar comments is indicated in parentheses. The approval rating of each methods is also shown.

Method	Approval rating	Positive comments	Negative comments
General	N/A	The scenarios were realistic (10)	
		I need increased training and practice to make weight estimation easier and more accurate (17)	
Broselow tape	2/32 (6%)	Broselow tape quick and easy to use (6)	Broselow tape not long enough (12)
			Broselow tape difficult to use in non-supine children (6)
EDD4C app	10/32 (31%)	Quick and easy to use (15)	Using scroller to insert length was tedious and slow (13)
		No need for calculations (8)	App dependent on Wi-Fi signal (5)
Mercy method	3/32 (9%)	Mercy method easy to use in sitting kids (4)	Mercy method difficult, complex and complicated (8)
		Mercy method accounts for habitus discrepancies (1)	Mercy method difficult in uncooperative patients (5)
		Actual, objective measurements used (2)	
PAWPER XL tape	17/32 (53%)	Quick and easy to use (25)	PAWPER XL tape easy in supine but difficult in sitting patients (6)

10.3 Epilogue

10.3.1 Major findings

This study was the first to evaluate the performance of weight estimation systems during simulated emergencies, with paediatric patients in different positions and with varying degrees of cooperation. It was also the first study to evaluate the use of a mobile phone weight estimation app.

Table 10-1 below summarises the major findings for each weight estimation method with regard to the effects of human factor errors, patient position (supine, lateral or seated) and degree of patient cooperation.

Table10-1 The influence of each factor on the accuracy of the weight estimation method. The effect size category is indicated in each block. An odds ratio of 1.6 was considered to indicate a small effect size, 3.5 a medium effect size and 6.5 a large effect size [12].

	Human factors	**Patient position**	**Patient cooperation**
Broselow tape	Medium	Large	Large
Mercy method	Large	Medium	Medium
EDD4C app	Small	Medium	None
PAWPER XL tape	Small	None	None

The surprising findings were that the Broselow tape was most vulnerable to both patient and human factor errors and that the Mercy method was affected by patient factors to a far lesser degree. The unsurprising, but important, findings were that human factor errors had a significant effect on weight estimation accuracy for all the systems that were evaluated.

The time taken for the weight estimations were similar for the Broselow and PAWPER XL tapes and slightly longer for the other methods. These times were essentially unaffected by patient factors in this study.

10.3.2 This study in the context of the thesis

This was the most unique study in this thesis as well as one of the most important. Another recent study on the usability of formulas to estimate weight in emergencies also highlighted the difficulty of using simple techniques under real-life circumstances

[4]. Studies like this can provide useful information to allow users to make an informed choice on the best weight estimation system to use in any clinical scenario.

In the context of this thesis, the study confirmed the effectiveness of both the PAWPER XL system and the Mercy method during realistic simulated emergency medical care. Both these methods have convincingly shown that they are two of the top three weight estimation methods available at present.

Other factors were now identified as having acquired greater significance: much more research is needed on how real-life problems can impact on weight estimation. This concept of real-life problems has a broader significance and is important not only for weight estimation systems that require measurements, but also methods such as parental estimates. For example, it is not known how often parents are actually available to provide a weight estimation at the time that it is needed. It is also not known whether these parental estimates are as accurate during the inevitable stress generated by the situation, as they are at other times.

Human factor errors had the greatest impact on the accuracy of the PAWPER XL tape and the Mercy method. A previous study by Abdel-Rahman and colleagues (the developers of the Mercy systems) showed very large differences between the accuracy of estimations using the Mercy tapes obtained by novice and trained users under non-emergency conditions [8]. These differences were of a similar magnitude to what we observed in this study. While the differences between the best and worst raters in our study were also less extreme with the PAWPER system, they were still significant. This again raises one of the common themes running through the thesis: training of users is essential to enable them to use weight estimation systems accurately. Paediatric patient safety may depend on training in weight estimation in the same way as it depends on training in resuscitation skills.

Another positive finding of this study is that it is possible to tentatively speculate about the PAWPER XL-MAC system from the findings on the Mercy method. Since both the PAWPER XL tape and the Mercy method, which share elements of methodology with the PAWPER XL-MAC system, were relatively unaffected by

patient position and cooperation, it is possible that the same will be true for the PAWPER XL-MAC system. This will need further research, however.

The significant findings in terms of accuracy and usability were that the PAWPER XL tape showed the highest accuracy and usability, the Mercy method was accurate but less usable while the EDD4C App was only slightly better than the Mercy method. The Broselow tape was shown to be neither accurate nor usable.

The EDD4C App showed promise, but further research and development will be required to optimise its functioning and establish the role of similar Apps in this field.

10.3.3 Link to the next section

The following section contains the summation of the work from this thesis: the final integrated discussion, the recommendations and suggestions for future research as well as the final conclusions.

10.3.4 References

[1] Lancaster L. Mission impossible — minimising the terror of paediatric resuscitation for staff in the ED. Australasian Emergency Nursing Journal. 2002;5(3):16-20.

[2] Marikar D, Varshneya K, Wahid A, Apakama O. Just too many things to remember? A survey of paediatric trainees' recall of Advanced Paediatric Life Support (APLS) weight estimation formulae. Archives of Disease in Childhood. 2013;98(11):921.

[3] Carley S. APLS weight estimation – don't do it (well almost never). 2013 [Accessed March 2018]. Available from: http://stemlynsblog.org/apls-estimation-formulas-do-not-safely-predict-weight-in-uk-children-st-emlyns/#comments

[4] Marlow RD, Wood DLB, Lyttle MD. Comparing the usability of paediatric weight estimation methods: a simulation study. Archives of Disease in Childhood. 2018.

[5] Bernius M, Thibodeau B, Jones A, Clothier B, Witting M. Prevention of pediatric drug calculation errors by prehospital care providers. Prehospital Emergency Care. 2008;12(4):486-94.

[6] Pentin J, Smith J. Drug calculations: are they safer with or without a calculator? British Journal of Nursing. 2006;15(14):778-781.

[7] Kaufmann J, Laschat M, Wappler F. Medication errors in pediatric emergencies: a systematic analysis. Deutsches Ärzteblatt International. 2012;109(38):609-16.

[8] Abdel-Rahman SM, Jacobsen R, Watts JL, Doyle SL, O'Malley DM, Hefner TD, et al. Comparative performance of pediatric weight estimation techniques: a human factor errors analysis. Pediatric Emergency Care. 2017;33(8):548-552.

[9] Luten R, Wears RL, Broselow J, Croskerry P, Joseph MM, Frush K. Managing the unique size-related issues of pediatric resuscitation: reducing cognitive load with resuscitation aids. Academic Emergency Medicine. 2002;9(8):840-7.

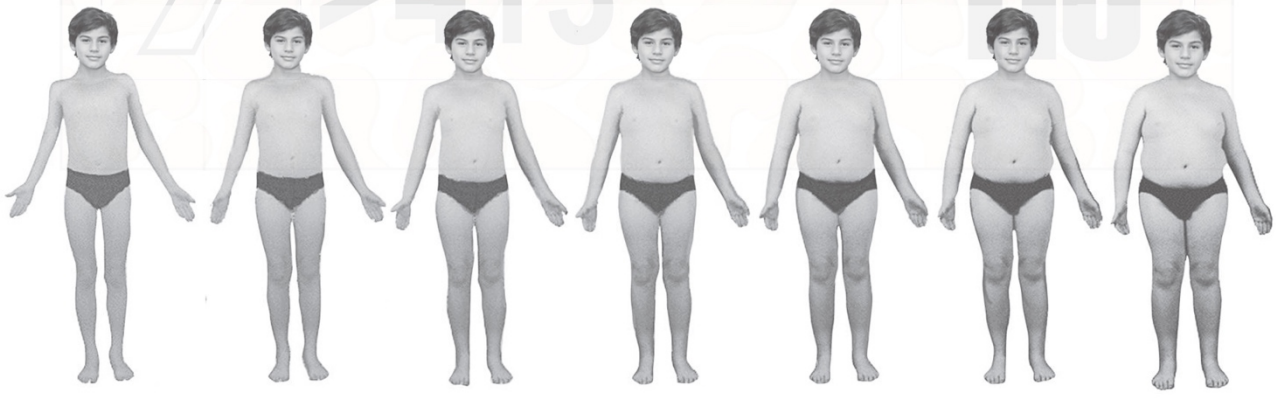
[10] Hohenhaus SM, Frush KS. Pediatric patient safety in resuscitation simulations: lessons learned from studying clinician behavior. Pediatric Emergency Care. 2005;21(1):71-75.

[11] Hohenhaus SM, Frush KS. Pediatric patient safety: common problems in the use of resuscitative aids for simplifying pediatric emergency care. Journal of Emergency Nursing. 2004;30(1):49-51.

[12] Chen HC. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. Communications in Statistics - Simulation and Computation 2010;39(4):860-864.

Section 4

Integrated discussion and conclusions



Chapter 11

Integrated discussion

Contents

Chapter 11 288

Integrated discussion 288

Contents 288

11.1 Introduction 290

11.1.1 A note on the Mercy method 290

11.1.2 The research objectives 291

11.2 The accuracy of weight estimation systems 291

11.2.1 The effect of weight estimation accuracy on outcome 291

11.2.2 A motivation for the need for accurate weight estimation 292

11.2.3 Standards required for weight estimation systems 293

11.2.4 The meta-analysis findings 294

11.2.5 Accuracy outcomes in low-and middle-income countries 295

11.2.6 Special circumstances: the accuracy of the Broselow tape 296

11.2.7 Special circumstances: the accuracy of age-formulas 297

11.2.8 Summary: the new standards in weight estimation 298

11.3 Refining the performance of the PAWPER system 298

11.3.1 Why was there a need for change? 298

11.3.2 The change in length 299

11.3.3 The additional habitus score categories 299

11.3.4 The estimation of ideal body weight 301

11.3.5 Categorising habitus: anthroposcopic methods 302

11.3.6 Categorising habitus: anthropometric methods 304

11.3.7 Did the PAWPER XL tape outperform the original PAWPER tape? 305

11.3.8 Summary: successful new models of the PAWPER system 305

11.4 The usability of weight estimation systems 306

11.4.1 The concept of usability 306

11.4.2 Which systems were the most usable? 307

11.4.3 Human factor errors 307

11.4.4 Patient factor errors 309

11.4.5 Weight estimation system factor errors 310

11.4.6 Summary: which systems were the most usable? 311

11.5 Integration of weight estimation and drug dosing 312
11.5.1 Why is integration important? 312
11.5.2 Is there evidence for the use of length-based tapes with pre-calculated doses? 312
11.5.3 Implications for the PAWPER system 314
11.5.4 Viable methods of integrating weight estimation and drug dosing 314
11.5.5 Summary: the co-dependence of weight estimation and drug dosing 315
11.6 Teaching and training 316
11.6.1 Why was training identified as important? 316
11.6.2 Evidence from other studies 316
11.6.3 The objectives of teaching and training 317
11.6.4 Methods of teaching and training 317
11.6.5 Summary: the importance of training 318
11.7 The need for mobile phone technology 318
11.7.1 How accurate and usable was the mobile phone application? 318
11.7.2 The need for mobile phone applications 318
11.7.3 Summary: the role of mobile phone applications 319
11.8 Summary of key findings 319
11.9 References 322

11.1 Introduction

For a healthcare provider, knowing the weight of a child during an emergency is critically important [1]. Having confidence that the weight is accurate, might make the difference between the healthcare provider administering or omitting a necessary dose of an essential medication [2]. This type of research is important to ensure that an accurate weight is obtainable during a paediatric emergency and to ensure that healthcare providers have the confidence to use a system in real-life emergencies.

The research questions that existed at the beginning of this research evolved over the course of the projects. This was primarily because of the publication of new information that was relevant to paediatric weight estimation in general as well as to the PAWPER tape specifically. The new information that became available was the publication of the first Mercy method study as well as data from three international studies on the PAWPER tape [3-6]. The findings of these studies mandated the inclusion of an evaluation of the Mercy method into this research. They also provided additional guidance for the already-recognised necessity to modify the PAWPER tape to improve its accuracy in populations outside of South Africa.

11.1.1 A note on the Mercy method

At the start of this research it was not clear which weight estimation systems were the most globally accurate. During the course of the research, if a system had been identified that was more appropriate for emergency use, clearly superior in accuracy and usability, then this system would have been pursued in further research, in addition to further exploratory and refinement work with the PAWPER systems. In addition to weight estimation accuracy, the other factors that would have been considered in deciding between systems was their appropriateness for emergency use and their ability to accurately predict IBW.

Although there was no method that was clearly superior to the PAWPER tape, the Mercy method was very similar to the PAWPER tape in terms of accuracy. The initial Mercy method study was published in a small non-indexed journal in August 2012. Other articles followed from the Mercy team between 2013 and 2015. After the initial

systematic reviews, the Mercy method was included in several other studies that were conducted subsequently: four studies not included in this thesis and two studies that were part of the doctoral research. These studies provided additional clarity on the performance of the Mercy method as well as allowing a direct comparison with the accuracy of the PAWPER tape.

Additional research is required for both the PAWPER systems and the Mercy method, but this was beyond the scope of this thesis, however.

11.1.2 The research objectives

The objectives of the research presented in this thesis can be summarised as follows:

- to critically assess the accuracy of existing weight estimation systems
- to determine whether the Broselow tape was an appropriate exemplar for the integration of a weight estimation tool and a drug dosing instrument
- to develop and critically evaluate refinements to the PAWPER system
- to assess the usability and reliability of several different weight estimation systems when used under adverse conditions.

The outcomes of the studies in this thesis will be discussed in terms of how they have addressed the objectives, the problem statements and the motivation for the research.

11.2 The accuracy of weight estimation systems

11.2.1 The effect of weight estimation accuracy on outcome

Is it actually important to acquire an accurate estimation of weight during the management of medical emergencies in children? In the findings of the systematic reviews in this thesis, there was no direct evidence to answer this question and expert opinion in the literature was divided on the matter. Nonetheless, avoiding harm from medication errors should be a high priority and the accuracy of weight estimation plays an important role in this regard [7]. During emergency care, if actual weight is not known, an estimation of weight is unavoidable in order to permit drug

therapy. The highest standard of care (the most accurate estimation practically achievable) is therefore essential. A prominent weight estimation expert's statement from 2007 is apt: "[…] to continue to use a demonstrably poor weight estimate cannot be considered good medical practice" [8].

11.2.2 A motivation for the need for accurate weight estimation

Failure to obtain an early and accurate weight estimate may be a major impediment to resuscitative care for two reasons: firstly, the process of obtaining a weight estimation may cause uncertainty and delay which will prevent the timely administration of necessary medications as well as other weight-dependent interventions (e.g. defibrillation); secondly, inaccurate weight estimation may result in dosage errors, with the resultant potential for patient harm [9]. Failure to obtain a timeous, accurate and appropriate weight estimation must be considered a medical error [10]. Ultimately, the accuracy of the weight estimation may determine the efficacy and safety of resuscitative drug treatment and failure to achieve accurate dosing may lead to increased mortality, especially in obese or underweight children who are at highest risk of inaccurate weight estimations [11].

In a recent study on paediatric inpatient medication errors, every weight estimation error >10% had a 33% chance of producing a medication dose error [12]. A recent systematic review concluded that nearly 50% of patients who receive a weight estimation error will have significant medication errors with a variable potential to cause harm, depending on the drugs prescribed [13]. Drugs used in acute medical emergencies often have a higher potential for harm than those used more routinely e.g. antiarrhythmics vs antibiotics [14].

Compounded and cumulative errors which occur during the dilution, preparation and administration of intravenous medications during emergencies may already produce an error of >20% in a significant number of drug doses [15, 16]. If these doses are based on incorrect weight estimations then the final dose error will be substantially higher.

These issues can be illustrated by continuing with the example of a child who presents with status epilepticus, mentioned in Chapter 2. The management of status epilepticus is time-critical as the longer the seizures last, the more resistant they become to treatment [17]. Any delay to treatment, which may occur as a result of dosing problems (including weight estimation), can increase mortality. Previous studies have shown that benzodiazepine dosing has been outside usual dosing guidelines in 23% of children with status epilepticus, as a result of weight estimation and/or dose calculation errors [18]. While excess dosing is common and can contribute to respiratory depression and increase the need for PICU treatment, inadequate dosing is also frequent and associated with poorer outcomes [18, 19]. Titrated doses are not appropriate in this scenario: it is essential to get the doses correct the first time. This also means that an accurate weight estimation is indispensable.

The conclusion that can be drawn from the literature, is that an accurate weight estimation, ideally with <10% error, is required in order to ensure patient safety from dose-related medication errors [12].

11.2.3 Standards required for weight estimation systems

In the qualitative part of the systematic review in Chapter 4, 150 articles were evaluated to establish whether there was any evidence about an acceptable standard of accuracy for a weight estimation system. However, virtually no evidence was found on benchmarking standards and very few articles even suggested an appropriate accuracy of weight estimation. There was therefore no evidence to justify any particular minimum standard of accuracy for weight estimation systems. The only previously published proposal of an acceptable outcome target recommended that more than 95% of weight estimations should fall within 20% of actual weight (PW20>95%) [20]. This proposed criterion was modified to provide a more specific benchmark for acceptable accuracy: more than 70% of weight estimations should fall within 10% of actual weight (PW10>70%) and more than 95% of weight estimations should fall within 20% of actual weight (PW20>95%). While this standard might be regarded by some experts as overly stringent, weight estimation accuracy is an important factor in patient safety and the highest standards should be embraced [21].

Many authors seem resigned to accepting that inaccurate weight estimation is inevitable: this needs to be challenged and reconsidered [22].

11.2.4 The meta-analysis findings

In the only previous systematic review on weight estimation, published in 2016, Young and colleagues concluded that the most accurate weight estimation systems worldwide were the Mercy method, the PAWPER tape and parental estimates of weight [23]. These findings were replicated in both the global meta-analysis from this thesis and a meta-analysis focused on studies from low- and middle-income countries [24, 25]. After analysing pooled data from nearly 100 studies, the following key findings were identified:

- No one-dimensional weight estimation system, whether based on age or length, achieved acceptable accuracy. These systems produced substantial underestimation of weight in obese children and considerable overestimation of weight in underweight children. These errors have the potential to cause patient harm.
- Parental estimates were in the top five most accurate methods in all the analyses. Although parental estimates of their children's weight may be accurate, as was found in this study as well as a previous systematic review, the accuracy of these estimates has not been fully validated in low- and middle-income countries [26]. There are some preconditions that have been recognised to ensure that an accurate parental weight estimation can be identified: the parent must be present, able and willing to provide a weight estimate; the parent must be the regular caregiver for the child; and the child must have been weighed in the parent's presence within the last six weeks [27]. There is also some doubt as to whether parents estimate the weight of obese children accurately [28, 29]. In circumstances when parents are not available to provide an estimate of weight, a backup system is required.
- The two-dimensional, length- and habitus-based systems proved to be more accurate than the one-dimensional systems, with the Mercy method and the PAWPER tape proving to be the most accurate of all the evaluated weight estimation methods. They achieved the lowest bias, the highest precision and

an acceptable accuracy in a variety of populations from both high-income and low- and middle-income countries. This reflected a reasonably consistent pattern across studies, including those studies in which the two-dimensional systems had performed less well than expected, although still performing better than other methods [4, 5]. It would be appropriate that these systems be developed and researched further for clinical use in preference to one-dimensional systems.

One of the interesting findings from the subgroup analyses was that almost all the weight estimation systems were accurate in normal weight children. The dual length- and habitus-based systems had a substantially higher accuracy in underweight, overweight and obese children, however, which resulted in a greater overall accuracy in each study. The other important observation was that weight estimation systems can appear more accurate than they are when they are evaluated in samples with a high proportion of normal weight children. For this reason, subgroup analysis is important to evaluate accuracy in children with a habitus that deviates from the norm.

11.2.5 Accuracy outcomes in low-and middle-income countries

There were no formal systematic reviews or meta-analyses of weight estimation studies from low-and middle-income countries until the one conducted in association with this research [25]. This review and several recent studies have shown the following outcomes:

- Children in low- and middle-income countries are at risk of dangerous overestimation of weight using one-dimensional age- and length-based methods [30, 31]. Most age formulas as well as the Broselow tape were designed and derived from high-income country populations, so they will tend to overestimate weight in underweight children. The changes to the Broselow tape over the years have exacerbated this overestimation. Although these systems have not been shown to be accurate in any population, they pose a hazard of overdosing underweight children, which is potentially more directly harmful than underdosing.

- The dual length- and habitus-based systems, such as the Mercy method and the PAWPER tape did not show the same trends as the Broselow tape and were the most accurate systems for use in low- and middle-income country populations.
- There have been no studies on the accuracy of parental weight estimation in resource-limited environments [32]. The few studies from urban areas in middle-income countries may not be representative of poorer populations [33]. There is considerable doubt as to whether parental estimates would be as reliable as suggested in high-income countries.

This meta-analysis strongly supported the use of the Mercy method and the PAWPER tape as the current most accurate paediatric weight estimation systems in low- and middle-income countries.

11.2.6 Special circumstances: the accuracy of the Broselow tape

A specific discussion of the Broselow tape is useful because the Broselow tape is generally considered to be the "gold standard" for weight estimation in children. However, it has performed well in very few studies, which have generally been in populations from middle-income countries with a high prevalence of normal weight children [34, 35]. The authors of studies from high-income countries have, for many years, expressed concern that the Broselow tape underestimates the weight of children to such a degree as to impact on medical care [36]. On the other hand, more recent evidence from low- and middle-income countries has shown that the overestimation of weight by the newest Broselow tape 2011 edition A has reached dangerous levels [30, 31, 37]. A definitive answer to the role of the Broselow tape is therefore very important.

The meta-analysis data showed that the Broselow tape was significantly less accurate than both the Mercy method and the PAWPER tape when compared directly. The data were clear that the Broselow tape was not as accurate as newer methods of weight estimation. Its poor accuracy and the limited value of the pre-

calculated drug data on the tape mean that its use and its status as the “gold standard” method should be reconsidered (see sections 11.2.8 and 11.5).

One of the other considerations was that some advanced life support organisations regard a child’s TBW as essentially equivalent to Broselow tape weight (which is actually closer to IBW) from a drug dosing perspective [38]. This is no longer considered to be valid in underweight or obese children and this false perception also needs to change [39].

More recent information has shown that modifying the Broselow tape weight estimation based on body habitus might prove useful in underweight and obese children [40, 41]. This will prove to be an important consideration in future research but the question around the drug information remains unresolved.

11.2.7 Special circumstances: the accuracy of age-formulas

Since age-based weight estimation formulas are so widely taught, researched and used in clinical practice, it was also necessary to deal with them specifically [42]. The meta-analyses as well as the study which evaluated every previously published age-based formula, showed definitively that age-based weight estimation is very inaccurate in every population. Multiple previous studies have had similar findings and yet clinical practice (or teaching practice) has not changed [43, 44]. Even the recently published habitus-modified, age-based formulas, which showed potential in the descriptive study, were shown to be inaccurate [45]. The most-recent studies using sophisticated age-, habitus- and sex-based computer generated weight estimates have shown some promise, but remain to be validated [46].

The conclusion is that no age-based formulas should be used [47]. They are not accurate in any population and the need for memorisation and calculations increases the risk of error [48]. The only possible exception where age formulas might be acceptable is in young children (<5 years of age) in selected communities in which the formulas have been tested and found to be accurate [6]. As a general rule,

however, the use of age formulas should be abandoned: there are more accurate weight estimation systems available.

11.2.8 Summary: the new standards in weight estimation

Although the Broselow tape has long been considered the standard for paediatric weight estimation, both the PAWPER tape and the Mercy method have been shown to have a significantly higher global weight estimation accuracy [49]. Parental estimates of weight are also more accurate than the Broselow tape, if the conditions that determine the accuracy of parental estimations are met.

In obese children specifically, parental estimates might not be accurate. Furthermore, IBW might need to be estimated using the PAWPER tape or Broselow tape, in any event, if drug therapy is required [22].

The current standards of weight estimation must therefore be considered to be the Mercy method, the PAWPER tape and parental estimates (the latter with provisos). If the usability of the methods is considered (see section 11.4), the PAWPER tape could be considered to be both more accurate and more usable than the Mercy method and may be preferred during emergency care.

11.3 Refining the performance of the PAWPER system

11.3.1 Why was there a need for change?

Three factors were identified as limitations in the functioning of the original PAWPER tape from previous studies, as discussed previously in Chapter 2. Design changes to address these factors, plus a method to estimate IBW, were incorporated in the new PAWPER XL system (Table 11-1).

Table 11-1 Changes to the PAWPER tape system.

Original PAWPER tape limitation	PAWPER XL tape solution
The tape, at 153 cm in length, was too short to provide a weight estimation for a substantial number of children.	The tape length was increased to 180 cm which allowed for weight estimation in children up to the age of 16 years.
Could not provide accurate estimations for severely obese children.	Two additional habitus score categories were introduced for these children.
The assessment of habitus was subjective and vulnerable to error.	Two new objective methods of categorising habitus were devised and validated – using image-assisted habitus assessment and using MAC measurements to determine habitus.
The tape was not intended to estimate IBW in addition to TBW.	The tape was specifically designed to be able to estimate both TBW and IBW.

The effects of these changes, as they were evaluated in the research presented in this thesis, will be discussed in this section.

11.3.2 The change in length

As many as 40% of children that present to the Emergency Department are over the age of 10 years [50]. Of these children, more than 70% will be too tall for the original PAWPER tape to provide a weight estimation [51]. This creates a problem in how to obtain a weight estimation in these children, as another method of weight estimation would need to be used. The change in length of the PAWPER XL tape has addressed that limitation: the length of 180 cm is sufficient to provide a weight estimation for 95% of children up to the age of 18 years, according to NHANES data from the USA [52]. In the study included in Chapter 7, nearly 10% of the children were too tall for the original PAWPER tape but none were too tall for the PAWPER XL tape [53]. The accuracy of weight estimations in this subgroup of taller children was equivalent to that found in shorter children. Although the PAWPER XL tape needs further prospective validation in taller children, the initial data was promising.

11.3.3 The additional habitus score categories

The outcomes of obese children suffering a cardiac arrest are worse when compared with the outcomes of normal weight children. This may be because of incorrect drug dosing in these children resulting from inaccurate or incorrect weight estimation [11, 54]. An accurate estimation of weight is needed for all children in order to prescribe

medications correctly. Since the prevalence of obesity is high in many populations across the world, including many populations in low- and middle-income countries, a weight estimation system should ideally be able to accommodate underweight, normal weight and obese children [55].

The question of how many habitus score categories to include in the new tape was also considered. The evidence from body image assessment studies suggested that seven categories were appropriate to estimate body habitus with the necessary degree of precision [56]. Since the new PAWPER XL model was intended to be used with reference images (see below) to assist habitus assessment, the findings of these studies were considered helpful.

The evidence from the weight estimation literature about the optimum number of habitus categories was less clear. The first ever published dual length- and habitus-based weight estimation system, the Devised Weight Estimating Method (DWEM), made use of three habitus categories [57]. Two more recent studies compared the accuracy of weight estimation using either three or five habitus categories and found no significant difference between the two [58, 59]. Other studies have made use of six or seven categories with good results [46, 60]. The optimum number of categories is likely to vary between systems depending on the methodology, but still needs to be resolved with further research.

In the validation study for the PAWPER XL tape, which evaluated the accuracy of the new tape in obese children, nearly 20% of the children in the study population were obese [53]. The PAWPER XL tape estimated weight significantly more accurately than the original PAWPER tape in these children. While the accuracy of the new tape did not achieve the levels of acceptable outcome (PW10>70%) in the subgroup of severely obese children, it was substantially better than the original tape. The accuracy of the PAWPER XL tape needs to be evaluated more extensively in other populations with a high prevalence of obesity. If this finding is reproduced in other studies, then a recalibration of the upper habitus score predicted weights may be necessary. Nonetheless, the PAWPER XL tape successfully increased the accuracy

of weight estimation in obese children through the addition of two extra habitus score categories.

11.3.4 The estimation of ideal body weight

With the increased focus on accurate weight estimation in obese children, it also became necessary to consider the importance of drug dosing in these children. It has not yet become well-accepted to consider tailoring drug doses according to body habitus during emergencies because of a lack of evidence showing changes in outcome [38]. Some preliminary evidence has suggested that obese children may suffer worse outcomes as a result of errors in drug dosing, however [11, 54]. The uncertainty with how to dose obese children has also been identified as an impediment to good emergency care [54]. Given that patient safety is an imperative during medical care, further research into the practicality of variable drug scaling in obese children is essential.

There is now reasonable consensus on how drug doses should be scaled in obese children [61-65]. Lipophilic drugs should be scaled to estimated or measured TBW and hydrophilic drugs scaled to estimated or calculated IBW. Unfortunately, the use of a single “pharmacokinetic mass” as a weight descriptor or scalar for drug dosing in obese children is not physiologically feasible [66]. Fortunately, the extremely close relationship between IBW and recumbent length makes it easy to predict IBW using tape-based methods. Ideally, then, a weight-estimation system should be able to estimate *both* TBW and IBW accurately, to prevent medication errors and potentially poor outcomes, especially in obese children. The importance of using drug doses scaled according to theoretical considerations of pharmacokinetics in obese children, in terms of improved outcomes, has not yet been established [38]. This warrants additional investigation.

In the study in Chapter 7, both the PAWPER XL tape and the Broselow tape were shown to be able to estimate IBW with a high level of accuracy [53]. This was confirmed in a subsequent study in which the PAWPER XL tape achieved an accuracy of 100% of estimates within 10% of IBW [22]. The ability of the PAWPER

XL tape to estimate both TBW and IBW accurately and easily is of great potential benefit for the user who wishes to optimise emergency drug dosing in obese children. This is crucial because the consequences of recent recommendations on drug dose calculations in obese children are that *both* TBW and IBW are required for accurate drug dosing (of different classes of drugs) [62-65]. At this time, only the PAWPER XL tape can provide both these weight descriptors with minimal cognitive burden.

11.3.5 Categorising habitus: anthroposcopic methods

Errors in the subjective visual assessment of habitus, leading to a reduced accuracy of the original PAWPER tape, had two potential solutions. The first solution was to explore methods of improving and standardising the visual, or anthroposcopic, assessment of habitus. The second solution was to explore alternative, highly objective methods of assessing and quantifying habitus using anthropometric measurements (see Section 11.3.6).

Preliminary evidence on the use of the original PAWPER tape had shown that visual assessment of habitus was somewhat subjective and it was not performed well by some groups of healthcare providers [6]. One of the intentions of the doctoral research was to develop a checklist method of habitus assessment as well as a reference-image method of assessing habitus. This was accomplished and then evaluated in the study presented in Chapter 8, to determine which habitus assessment method produced the greatest accuracy of weight estimation when used with the PAWPER XL system.

In this study, in novices with no prior experience with the PAWPER system, both the checklist and reference-image systems were found to be more accurate than the visual gestalt system. Overall, the use of reference images to assess habitus was the most objective, most accurate, most cognitively neutral and most satisfactory to the participants. Both doctors and paramedics achieved high accuracy using this technique and paediatricians were no more accurate than other doctors or paramedics at habitus assessment. However, nurses were less accurate at correctly

scoring habitus. There was no convincing explanation for this, but it has important implications for clinical practice which need to be explored in future studies.

There is evidence from other fields, especially body image research, which has provided support to the concept of using reference images to assess habitus [56, 67, 68]. These images have been shown to be well correlated with BMI as an objective indicator of habitus [69]. Since images provide a method of standardising assessment in different parts of the world (where the “average” child might have a very different habitus in high-income countries when compared to low-income countries), this could prove useful in increasing the accuracy of the tape. This method was therefore incorporated into the preferred methodology for the use of the PAWPER XL tape.

Given the fact that the use of images or silhouettes has been shown to be useful to predict BMI, it is unsurprising that they were found to be useful in this application. The use of age-appropriate images has also been shown to increase the accuracy of use [70, 71]. Only one study, from New Zealand, has previously made use of reference images to assist with length-based weight estimation [72]. This study found that the use of images enabled a greater accuracy of weight estimation in overweight and obese children. This provides additional support to the concept of image-assisted habitus assessment, but the matter still needs further research.

One of the most interesting findings in this study was the differences in accuracy produced within the group of apparently homogeneous novices. Although training clearly has a role in any skill-development, these were novices with no previous exposure to this technique [73]. The implications of human factor errors in weight estimation techniques has still not been fully elucidated [74]. It would be interesting to research the effect of training on individuals with poor initial habitus assessment accuracy.

One of the limitations of any form of visual habitus assessment (including the use of figural images) can be attributed to a visual bias described by Weber’s law. As applied to habitus assessment, this means that it becomes increasingly more difficult

to categorise the size of children as their BMI increases [75]. It may require more training for users to become accurate at classifying obese and severely obese children, than it would for identifying underweight children. This also will require further investigation.

11.3.6 Categorising habitus: anthropometric methods

Although the use of body habitus reference images improves the objectivity of anthroposcopic habitus assessment, the use of anthropometric measurements could make this process much more objective. Since MAC has been well-established as being strongly associated with both body weight as well as habitus, and is easy to measure, this was an obvious choice to research [3, 76]. MAC measurements have previously been used to define both underweight as well as obese status in paediatric nutrition studies [76, 77]. It was likely then that MAC measurement ranges would be able to be used to discriminate between habitus score categories.

The preliminary study, reported in Chapter 9, showed a good weight estimation accuracy of the newly derived PAWPER XL-MAC system in three diverse populations with varying prevalence of underweight and obese children [78]. The system achieved an accuracy exceeding the minimum benchmark requirements (PW10>70%). This accuracy was maintained in all habitus subgroups except in severely obese children. The level of accuracy shown by PAWPER XL-MAC system was very similar to that of the PAWPER XL tape, when used in two of the South Africa validation datasets [53, 79].

An important difference between the PAWPER XL-MAC methodology and other methods which make use of MAC, was that the PAWPER system completely eliminated the need for calculations. Both the Mercy method (which uses humerus length and MAC) and the Wozniak method (which uses ulna or tibial length and MAC) include a step dependent on arithmetic calculations [3, 77]. This could lead to delay and error if used during emergency care.

The acquisition of a measurement of MAC appears to be easily learnt and reproducible without the need for specific anthropometric training [76]. It has been shown to be highly reproducible with a high inter-observer agreement [80]. As with any technique, however, some training is required in order to ensure correct usage and accurate outcomes [74, 81]. The increased human factor errors seen in the use of the Mercy method during simulated emergencies (see Chapter 10) may also occur with the use of the PAWPER XL-MAC method, which also relies on a measurement of MAC. The best methods of training and how much training will be required to ensure that the PAWPER XL-MAC system is used accurately under such conditions will still need to be determined.

This technique makes use of measurements which can be obtained very rapidly and the absence of calculations should allow the methodology to be used with minimal cognitive burden [82]. Although this methodology still needs to be tested in a prospective study, it has shown promise as a more objective variation of the PAWPER XL tape.

11.3.7 Did the PAWPER XL tape outperform the original PAWPER tape?

The performance of the PAWPER XL tape was found to be superior to that of the original PAWPER tape in the study reported in Chapter 7 [53]. Unsurprisingly, the biggest improvements were in obese children and children too tall for the original tape. The PAWPER XL tape was also significantly more accurate than the Broselow tape and the Mercy method. This preliminary evidence was encouraging, but the new tape needs further validation by different users and in different populations.

11.3.8 Summary: successful new models of the PAWPER system

The principal refinements of the PAWPER system involved an expanded capacity to estimate weight in taller and fatter children and to provide alternative methods of assessing habitus that were more objective than simple visual inspection. The studies included in this thesis confirmed that:

- the changes to the tape allowed for weight estimation in taller children and increased the weight estimation accuracy in the obese children

- the use of reference images to guide the assessment of habitus significantly improve the accuracy of that assessment
- MAC could be used in a simple model to predict habitus, thus doing away with any form of subjective assessment.

The findings confirmed that the changes to the PAWPER tape achieved the objectives of the studies. The changes have expanded its ability to estimate weight and have increased its objectivity, without increasing the complexity of its functioning. Although these were preliminary studies, the findings were favourable. Further research will be required to confirm these findings and provide information that may further increase the accuracy of the system. The quest to minimise errors and ensure patient safety is an important priority.

11.4 The usability of weight estimation systems

11.4.1 The concept of usability

There are several different variables that can affect the usability of weight estimation systems, including human factors, patient factors and weight estimation system factors. It has been suggested previously that usability can be assessed by testing the accuracy, speed and user preference of weight estimation systems under the conditions in which they would be used in clinical practice [83]. Usability is a very important concept, because an accurate but impractical weight estimation system would not be useful in the clinical setting.

In the research for this thesis, both the accuracy and the usability of weight estimation systems were evaluated. The usability of selected weight estimation systems was assessed under the stressful conditions of simulated paediatric emergencies, with realistic clinical scenarios. There have been very few studies evaluating weight estimation systems in real or simulated emergency scenarios and no studies at all evaluating the accuracy of these systems in children in different positions. This is highly relevant, as emergency drug dose information is just as important for the child in acute respiratory distress (who might be in a seated position) as it is for a child in cardiac arrest (who would be in a supine position).

11.4.2 Which systems were the most usable?

This was the first time that the performance of weight estimation systems has been evaluated under simulated adverse, but realistic, conditions. The results were encouraging for the PAWPER XL tape and suggested that it would probably maintain the same degree of accuracy in a wide variety of clinical scenarios.

The accuracy of the weight estimation systems that were tested was similar to the accuracy reported in previous studies with cooperative children in ideal positions [23, 24]. The PAWPER XL tape exceeded the acceptable accuracy criteria proposed in this thesis, but the other systems did not. The accuracy of the PAWPER XL tape was unaffected by the position of the child (e.g. supine, lateral recumbent or recovery position or seated) or degree of cooperation (cooperative, passively uncooperative or actively uncooperative) in simulated medical emergencies. The performance of the EDD4C app was only affected by patient position but not cooperation. The Mercy method was affected by both patient position and cooperation, but to a lesser degree than the Broselow tape, which was substantially affected by each of these factors. Although patient factors played an important role in the accuracy outcomes, human factors were responsible for the biggest differences in accuracy outcomes for all the weight estimation systems.

The individual contributors to the usability, as reflected in the total error, human factor errors, patient factor errors and weight estimation system factors, will be discussed below.

11.4.3 Human factor errors

The most important determinants of the accuracy of the weight estimation systems were human factor errors [74]. Some participants achieved a near-perfect weight estimation accuracy with each of the systems (other than the Broselow tape) while others were exceptionally inaccurate. This has been reported previously in studies evaluating human factor errors [84-87]. There were no clear predictors of which participants would be more and which would be less accurate. Whether the human

factor errors were as a result of cognitive stress or training-related issues was not determined. This would be an important additional avenue of future investigation.

The difference in weight estimation accuracy achieved by the most and least accurate participants was instructive. The high levels of accuracy attained by the best participants showed the inherent potential of the systems' performance and confirmed the usability of the systems. The low levels of accuracy achieved by the worst participants exposed the impact of human factor errors on the weight estimation process. While the PAWPER XL system was the most resilient of the weight estimation systems evaluated, it was not immune to error when used by novices. This was vital evidence which suggested that sufficient training and practice is required to allow for the accurate use of weight estimation systems. The use of protocols and algorithms to optimise training and clinical management may be of value. This requires further investigation, as it has important consequences for how weight estimation should be taught and practiced.

The EDD4C app that was developed for the study performed well, although some errors and limitations unique to high-technology systems were detected, as has been noted previously with other systems [88, 89]. The slightly reduced accuracy in taller and non-supine children suggested that the use of the measuring tape may have been sub-optimal by some users. The use of the app needs further research and development to evaluate and address these limitations.

The Mercy method was the most vulnerable to human factor errors, with the greatest disparity between the best and worst users. It also had the lowest accuracy of all the systems amongst the worst users. This was probably as a result of the more complex methodology. Since the potential accuracy of this method was high, with the best users, it is likely that its usability would be significantly improved with training and practice. In other studies on the Mercy method – not during simulated emergencies – human factor error have been common when the method was used by undertrained individuals [74, 90]. Despite the fact that the developers of the Mercy method have never specifically promoted this method for use during emergencies, it is likely that it

will prove suitable in this role if used by individuals who have been appropriately trained.

The Broselow tape, despite being the simplest of the methods evaluated, was also susceptible to human factor errors. The interesting fact about this is that even simple systems can be vulnerable to errors and require training for accurate use [84].

User preference is both an important indicator of the usability of a weight estimation system. In the study in this thesis, the participants favoured the PAWPER XL tape system and the EDD4C app over the other methods because of ease-of-use and their perceived accuracy. This perception by users is important, as a system that is considered to be difficult to use or inaccurate is not likely to be easily adopted by clinicians.

There have been very few other similar studies with which to compare the findings from this article. The findings in those studies have all identified human factors as a major contributing factor to weight estimation errors and have all recommended training as the most important measure to address the problem [83, 85, 91].

Further research is required to identify methods of training which would allow all users to be accurate with the most accurate weight estimation systems such as the PAWPER XL tape, mobile phone applications (including the EDD4C app) and the Mercy method.

11.4.4 Patient factor errors

This study was the first to evaluate the usability of weight estimation systems under adverse conditions in terms of patient positioning and cooperation. This study was useful because these are the conditions under which the systems will actually be used in real-life scenarios. The resilience of the two-dimensional weight estimation systems, especially the PAWPER XL tape, to adverse patient factors was promising. Neither patient position nor degree of cooperation affected the accuracy of the tape. These findings strongly support the conclusion that the tape was both accurate and

usable. The EDD4C app was slightly less accurate in non-supine children, which was difficult to explain. It is possible that a combination of human, patient and method factor errors caused this inaccuracy (e.g. it was more difficult to accurately measure length using the measuring tape in taller, non-supine children). This requires further research to provide more clarity. The Mercy method was slightly less accurate in supine and uncooperative children. This was not surprising, given the nature of the measurements required for the methodology. The effect sizes were small, however, which does not rule out this method as a viable weight estimation system for use during emergency care. The Broselow tape was profoundly impacted by position and cooperation, perhaps because its large size made it difficult to wield in non-supine and uncooperative patients.

It was obvious that patient factors can play a significant role in the usability and accuracy of weight estimation systems. The future development of weight estimation methodology should include evaluation of these factors before they are considered for clinical use.

11.4.5 Weight estimation system factor errors

While the findings in this study confirmed the importance of human and patient factors in determining the accuracy and usability of weight estimation systems, it is likely that the equipment itself may contribute to the overall performance. This has been demonstrated previously with the 3D Mercy tape which has a significantly lower accuracy than the 2D tape because of the apparatus itself [74, 92]. Although this is partly a human factor error, it might be a factor that cannot be improved through training and practice.

There were two particular equipment-related factors that may have influenced the weight estimation process in this study. The first was the physical size of the Broselow tape. There have been no previous studies evaluating the effects of the size of weight estimation tapes on weight estimation accuracy during emergencies. Preliminary evidence from the use of the Broselow tape during simulated emergencies in this thesis suggested that the width of the tape may have made it

more difficult to manipulate in non-supine and in uncooperative children [93]. This was the most likely explanation for the discrepancies noted in the study. This finding needs to be assessed in future studies. If this is confirmed, then it suggests that other large-format tapes should be carefully evaluated in similar scenarios. The Sandell Smart tape, which is wider than the Broselow tape might be even more vulnerable to this error when used in non-supine or uncooperative children [94].

The second equipment-related error factor was related to the EDD4C app. This app, which was developed in this project, proved to be as accurate as the PAWPER XL tape system with respect to weight estimation accuracy. However, there was a different set of problems identified associated with using an app when compared to low-tech equivalents. These included vulnerabilities to interrupted Wi-Fi access, accidental closing of the app, user unfamiliarity with the use of electronic devices and other minor problems. These and other new technology-associated problems will need to be evaluated in future studies as they could potentially play an important role in the effectiveness (or not) of mobile phone applications.

11.4.6 Summary: which systems were the most usable?

There were four main learning points from this study in terms of the overall context of the thesis. The first point was that the PAWPER XL tape methodology proved to be resilient, accurate, usable in adverse circumstances and favoured by the participants. This was important because it confirmed that the slightly more complex two-step methodology of the PAWPER XL tape is viable, appropriate and accurate for emergency use. The second point was that the Mercy methodology was somewhat less resilient and usable than the PAWPER XL system. This suggests that the PAWPER XL tape may be more appropriate for emergency use than the Mercy method. This difference in usability may be less significant when the systems are used by highly-trained individuals, however. The fact that the Mercy method did show a large disparity between users may suggest that the PAWPER XL-MAC system might also be more difficult to use in adverse settings. Although this is speculation, it will be important to evaluate this in future studies. The third significant point was that the mobile phone application performed reasonably well. The ability of an app to

integrate weight estimation with drug dosing creates a potentially extremely powerful tool. The app will need to be developed further and subjected to additional rigorous testing. The final important point, alluded to above, is the need for training. It is clear that establishing and maintaining skills through training and practice is essential. Although weight estimation is a small part of providing emergency care, it is one of the skills that needs practice to ensure patient safety. This will be discussed further in section 11.6.

11.5 Integration of weight estimation and drug dosing

11.5.1 Why is integration important?

The major purpose of weight estimation during paediatric emergencies is to allow weight-based dosing of essential medications. The effective integration between the weight estimate and the drug dosing process is vital to ensure that an accurate weight estimation is translated into an accurate drug dose.

This process is more dependent on the drug dosing guide than the weight estimation system, but since the Broselow tape and other length-based tapes with pre-calculated drug doses are recommended by major advanced life support guidelines, it was important to discuss this concept and its implications for other methods.

11.5.2 Is there evidence for the use of length-based tapes with pre-calculated doses?

The main reason to explore the integration between weight estimation and drug dosing was to establish the validity of pre-calculated drug doses printed on some length-based weight estimation tapes. Some of the major international guidelines recommend the use these devices but it was not clear that the available evidence supported this [38]. Given the large number of these types of tapes that are currently available, it was important to review the existing evidence.

There are at least four tape systems that have drug dose information printed on the tapes themselves: the Broselow tape (and one version of the Pediatape), the Sandell

tape, the Handtevy tape and the *Padiatrisches Notfalllineal*. None of the tapes have been shown to estimate weight well, however [24, 25, 95-97]. Of these tapes, only the Broselow tape has been studied with respect to drug dosing accuracy in simulation studies of paediatric emergencies. One of the mooted advantages of the Broselow tape, in particular, is the availability of drug dose information on the tape itself [98]. The systematic review in Chapter 6 revealed that there was no evidence that the drug information on the Broselow tape (the resuscitation aid component of the tape) conferred any benefit during the emergency treatment of children [97]. The evidence was very sparse, but in every study in which the drug dosing accuracy of the unassisted Broselow tape was compared against another method, the other method was more accurate [16, 99-101]. In every study in which the Broselow tape proved superior to "standard" care, the Broselow tape was supplemented by additional reference material. This material removed the need for dilution and administration calculations by the user and therefore also nullified the need for drug dose information to be printed on the tape. The Broselow tape did not contain sufficient drug dosing information to function as a comprehensive, standalone resuscitation aid and there was insufficient evidence to show any benefit to the partial information that it presented.

These findings strongly suggest that there is no benefit to design future tapes with drug dosing information printed on the tape unless this information is sufficiently detailed to eliminate the need for supplementary resources. However, the physical size of the tape still remains a limitation in this regard [102]. It is noteworthy that more recently-designed tapes have followed the trend of avoiding drug dosing information on the tape itself for the reasons noted above [95].

The use of the Broselow tape, and other similar length-based tapes, should therefore be reconsidered. The guidelines recommending the use of these devices should also be revised.

11.5.3 Implications for the PAWPER system

With respect to the potential future design of the PAWPER system, the evidence was helpful to show that the PAWPER tape should be used to estimate weight only and that supplementary systems should be used to guide drug dosing [1].

11.5.4 Viable methods of integrating weight estimation and drug dosing

There are several ways in which weight estimation systems can be integrated with drug dosing systems:

- a companion aid in the form of a booklet or card
- a companion mobile phone application
- an integrated weight estimation and drug dosing mobile phone application
- novel colour-coded methods of administering medications without calculating doses.

The PAWPER tape system, the Kinder-Sicher system and the Paulino System all have companion guides to provide comprehensive information required to eliminate all decision-making and calculations in the drug dosing process [1, 95]. It is not widely known, but the Broselow tape also has a companion guide to simplify dosing calculations [103]. This guide is expensive (US$ 210), has physically large dimensions and is not available outside the USA. The PAWPER companion aid was evaluated in a recent simulation study in which it reduced dosing and cognitive errors when compared with the Broselow tape (unpublished data). These and similar other systems need to be further evaluated in simulated and real paediatric emergencies to confirm whether they function accurately and integrate efficiently with dual length- and habitus-based weight estimation systems.

There are several paediatric drug dosing applications available for Apple and Android devices on the iStore and the PlayStore. None of these applications have been studied in real or simulated emergencies to date. This would be important to establish their potential roles in addition to evaluating how well they integrate with the most accurate weight estimation systems.

The EDD4C application, developed as part of this research, is an integrated weight estimation and drug dosing application. It is the only one of its type that was identified through a search of the iStore and the PlayStore. It was evaluated in a simulation study with very encouraging results – it was more accurate than the Broselow tape system (unpublished data). Since there are substantial advantages to an integrated system such as this, including individualising dose scalars in obese children, this is a very important area for future research. The systems will need to be customised according to local needs and tested in real-life or appropriately simulated conditions.

The use of colour-coded syringes or other non-calculating methods of delivering drug doses has been shown to be more accurate than the Broselow tape system in several simulation studies [100, 101, 104, 105]. These systems have not been fully evaluated in terms of the accuracy of the final drug delivery, but rather only in the reduction of errors related to cognitive and calculation dosing errors in the standard treatment groups. They also need to be evaluated for accuracy against children's actual weight. Although these systems are a promising possibility for future use, further research will be required.

11.5.5 Summary: the co-dependence of weight estimation and drug dosing

There were three main concepts that emerged on the integration of weight estimation systems and drug dosing systems. Firstly, inaccurate weight estimation by the Broselow tape and the inadequacy of its drug dosing system suggest that its role in the resuscitation of children needs to be reconsidered. The international guidelines on the value of length-based tapes with pre-calculated drug doses need to likewise be revisited and revised. Secondly, this data was useful to clarify that the PAWPER system should not be changed to include drug doses on the tape. Thirdly, the practical integration of weight estimation and drug dosing systems needs further research to establish how it can best be achieved with the lowest cognitive demands and the greatest accuracy.

11.6 Teaching and training

11.6.1 Why was training identified as important?

Although a discussion of teaching and training was not one of the objectives of the thesis, it emerged as an important factor following the findings in two of the studies (Chapters 8 and 10). The first study showed that there was a major difference in the ability of different individuals to assess habitus correctly [106]. While some individuals were consistently exceptionally accurate, others were consistently inaccurate. There was no clear evidence or explanation for this phenomenon. The second study showed that certain individuals could achieve a high level of accuracy with different weight estimation methods used under adverse conditions in simulated emergencies. Other individuals performed poorly, also with no clear explanation about why this should be the case [107]. Since the level of experience was controlled in these studies, at least as much as was possible, the most likely explanation is that differing levels of aptitude and skill caused the disparities. This was a strong indicator that appropriate education and training, ideally with proof-of-learning outcomes, would be essential to ensure that all users could achieve an appropriate level of accuracy.

11.6.2 Evidence from other studies

There is strong evidence that the Broselow tape is prone to incorrect use when used by individuals with inadequate formal training with the tape [81, 84, 85, 108, 109]. There is also good evidence on the effectiveness of formal training or computer-based training in overcoming these vulnerabilities [86, 87, 110]. A study on human factor errors which evaluated several methods of weight estimation, including the Broselow tape and the Mercy tapes, showed a substantial difference in accuracy between well-trained and untrained users [74]. These studies showed that procedural errors are common with even the simplest weight estimation system, but that formalised training can be used to increase the accuracy of these systems by ordinary users.

11.6.3 The objectives of teaching and training

The objectives of training should be simple: to educate healthcare providers about the appropriate systems to use and the importance thereof; to impart the knowledge of how these systems work and when to use them in resuscitative care; to provide instruction and opportunity to practice with the system; and to provide training with the systems both in routine and emergency care situations.

Another important consideration would be to ensure that healthcare providers are adequately educated and trained in the entire process that they would have to follow in real-life practice: weight estimation, the use of a resuscitation aid for drug dose determination as well as the medication preparation and administration.

11.6.4 Methods of teaching and training

There is no published evidence to guide the methods of teaching and training in weight estimation. Advanced life support short courses would be an ideal environment to provide much of the initial education, familiarisation and practice with weight estimation systems, while ongoing practice could be conducted in the workplace. It would be important to include training during simulated emergencies, to induce in the users some degree of immunisation against stress as well as to ensure familiarity with the use of the weight estimation system in near real-life situations.

Some hospitals require that every emergency intubation is performed with a bougie to ensure that users are well-practised with this device when they need it for a difficult airway. In the same way, a weight estimation method could be used on every paediatric patient that presents to the Emergency Department – this would provide practice with the method and could increase accuracy because the user would get immediate feedback on their technique from the actual measured weight.

Further research is required to establish the type and frequency of training that would provide the optimum results.

11.6.5 Summary: the importance of training

Training and preparation form one of the cornerstones of Emergency Medicine. It was clear from the findings of the studies in this thesis, as well as findings from previous studies, that human factor errors have a significant influence on the accuracy of weight estimation systems. These errors occurred in both simple and more complex systems. Previous studies have shown that education and training is essential to reduce these errors. Formalised training should be provided for all weight estimation systems that are used in clinical practice. An algorithm could be used to aid both the use of and the training for weight estimation. It should be expected that more complex models such as the PAWPER XL-MAC systems may require more in-depth training to ensure accurate and consistent use. This will need to be evaluated in further research.

11.7 The need for mobile phone technology

11.7.1 How accurate and usable was the mobile phone application?

There were three important findings in the study reported in Chapter 10 with respect to mobile phone technology. The first was that the EDD4C app functioned well as a weight estimation device and was superior to the Broselow tape and Mercy method in terms of accuracy. The second finding was that the app was well-liked by the participants, who favoured it over all other methods except the PAWPER XL tape. The third important finding was that with technology came a different set of problems and challenges that need to be researched and addressed.

11.7.2 The need for mobile phone applications

There are very few apps available to provide information on weight estimation, drug dosing, or both. Part of the reason for this is regional or national differences in drug formulations and legislation governing the use of technology to provide dosing information.

There is little doubt that these applications can make the best use of the black box concept, which is important to minimise the cognitive burden generated by the process. The user would enter simple information and the application would automatically calculate the weight, determine the best dose scalar for each drug for each patient (e.g. TBW or IBW in obese children). In this way, the use of these applications would allow for the best integration possible between weight estimation and drug dosing guides. Furthermore, smartphones are generally available and affordable in most parts of the world, even in resource-limited environments, so this concept has true global potential [111].

11.7.3 Summary: the role of mobile phone applications

Since the role of drug information on length-based tapes has been cast into some doubt through the findings presented in this thesis, there is a need to find the best possible replacement solution. While paper-based systems must be developed, the potential value of mobile phone applications is far higher. The potential for mobile phone applications to be accurate and usable in weight estimation was established in Chapter 10 and the potential exists to link weight estimation and drug information in a single application. Although this was very preliminary work, and very limited research has been published on the use of applications in any aspect of Emergency Medicine, this topic is a definite priority for follow-up research.

11.8 Summary of key findings

The overall findings in this thesis can be summarised as follows:

- Age- and one-dimensional length-based weight estimation systems are not sufficiently accurate to ensure that patient harm is avoided when calculating medication doses from the weight estimates
- The most accurate existing weight estimation systems are the Mercy method and the PAWPER tape. Parental estimates are also amongst the most accurate methods, but have not been researched sufficiently, especially in low- and middle-income countries.

- There is no evidence that partial drug information contained on a length-based weight estimation tape is of any benefit to patient or healthcare provider. This was useful because it confirmed the validity of the PAWPER tape model: the tape is used to estimate weight, while an additional resource can be used to provide comprehensive drug dosing information.
- The PAWPER XL tape was developed as a response to the need for a longer tape and for a need to estimate weight better in obese children. It was more accurate than the original PAWPER tape and could provide weight estimations to an extended group of children: those up to 180 cm in length as well as obese and severely obese children.
- The PAWPER XL tape and the PAWPER XL-MAC system could predict both total and ideal body weight accurately. This is useful as it opens up the opportunity for optimising drug dosing in obese children, as IBW is easily and rapidly obtainable during emergencies.
- The use of body habitus reference images is more accurate than the original gestalt system for the PAWPER tape. These images should therefore be used to assist users in the assessment of habitus. The accuracy outcomes when using figural reference images was higher than that with a checklist method and with gestalt visual estimation.
- Adequate training is required to enable users to be accurate when assessing habitus for the PAWPER XL weight estimation system.
- The use of MAC with the PAWPER XL-MAC system was accurate and objective. It performed well in all habitus types except severely obese children.
- The PAWPER XL tape and the EDD4C app performed well in simulated paediatric emergencies. They remained accurate in children in the clinically realistic positions that might be found during emergency care. They were also equally accurate in cooperative and uncooperative children.
- The effect of human factor errors on the accuracy of the PAWPER XL tape was found to be important, both in the assessment of habitus as well as the use of the tape during emergencies. Some users achieved high levels of accuracy and some were less proficient.
- Figure 11-1 provides a graphic summary of the major findings in this thesis.

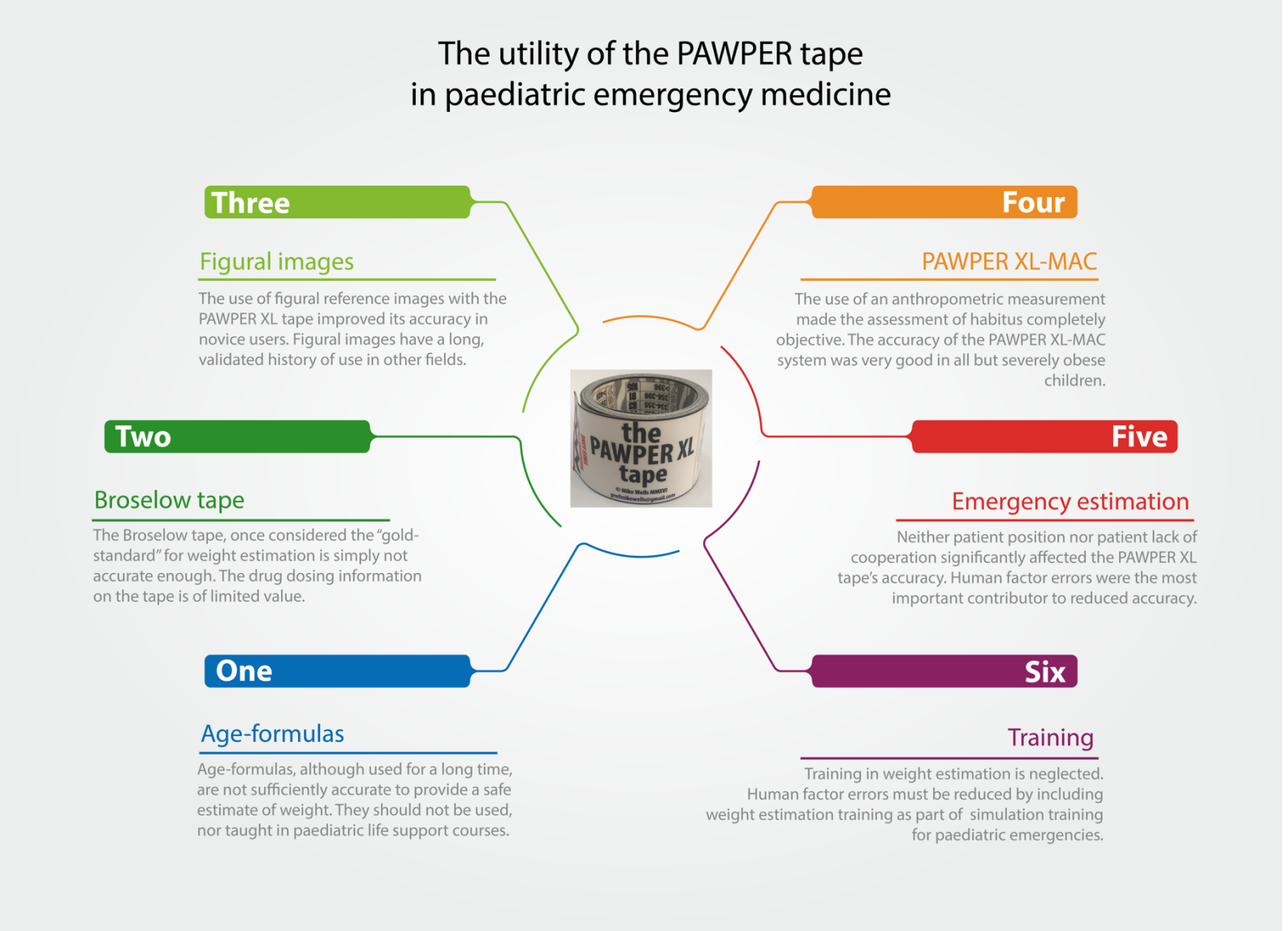


Figure 11-1 A summary of the major findings from this research.

11.9 References

[1] Wells M, Goldstein L, Botha M. Emergency drug dosing in children. Edinburgh: Churchill Livingstone Elsevier; 2012.

[2] Wells M, Coovadia A, Kramer E, Goldstein L. The PAWPER tape: A new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. Resuscitation. 2013;84(2):227-232.

[3] Abdel-Rahman SM, Ridge AL. An improved pediatric weight estimation strategy. Open Medical Devices Journal. 2012;4:87-97.

[4] Garcia CM, Meltzer JA, Chan KN, Cunningham SJ. A validation study of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape-a new weight estimation tool. Journal of Pediatrics. 2015;167(1):173-177.e1.

[5] Chavez H, Peterson R, Lo K, Arel M. Weight estimation in an inner city Pediatric Emergency Department: the effect of obesity. American Journal of Emergency Medicine. 2015;33(10):1364-1367.

[6] O'Leary F, John-Denny B, McGarvey K, Hann A, Pegiazoglou I, Peat J. Estimating the weight of ethnically diverse children attending an Australian emergency department: a prospective, blinded, comparison of age-based and length-based tools including Mercy, PAWPER and Broselow. Archives of Disease in Childhood. 2017;102(1):46-52.

[7] Lesar TS, Mitchell A, Sommo P. Medication safety in critically ill children. Clinical Pediatric Emergency Medicine. 2006;7:215-225.

[8] Luscombe M, Owens B. Weight estimation in resuscitation: is the current formula still valid? Archives of Disease in Childhood. 2007;92(5):412-415.

[9] Porter E, Barcega B, Kim TY. Analysis of medication errors in simulated pediatric resuscitation by residents. Western Journal of Emergency Medicine. 2014;15(4):486-90.

[10] Kaufmann J, Laschat M, Wappler F. Medication errors in pediatric emergencies: a systematic analysis. Deutsches Ärzteblatt International. 2012;109(38):609-16.

[11] Srinivasan V, Nadkarni VM, Helfaer MA, Carey SM, Berg RA, American Heart Association National Registry of Cardiopulmonary Resuscitation I. Childhood

obesity and survival after in-hospital pediatric cardiopulmonary resuscitation. Pediatrics. 2010;125(3):e481-8.

[12] Hirata KM, Kang AH, Ramirez GV, Kimata C, Yamamoto LG. Pediatric weight errors and resultant medication dosing errors in the Emergency Department. Pediatric Emergency Care. 2017.

[13] Pan SD, Zhu LL, Chen M, Xia P, Zhou Q. Weight-based dosing in medication use: what should we know? Patient Prefer Adherence. 2016;10:549-60.

[14] Lammers R, Byrwa M, Fales W. Root causes of errors in a simulated prehospital pediatric emergency. Academic Emergency Medicine. 2012;19(1):37-47.

[15] Kozer E, Seto W, Verjee Z, Parshuram C, Khattak S, Koren G, et al. Prospective observational study on the incidence of medication errors during simulated resuscitation in a paediatric emergency department. British Medical Journal. 2004;329:1321-1325.

[16] Feleke R, Kalynych CJ, Lundblom B, Wears R, Luten R, Kling D. Color coded medication safety system reduces community pediatric emergency nursing medication errors. Journal of Patient Safety. 2009;5(2):79-85.

[17] Sanchez Fernandez I, Gainza-Lein M, Abend NS, Anderson AE, Arya R, Brenton JN, et al. Factors associated with treatment delays in pediatric refractory convulsive status epilepticus. Neurology. 2018;90(19):e1692-e1701.

[18] Abend NS, Loddenkemper T. Management of pediatric status epilepticus. Curr Treat Options Neurol. 2014;16(7):301.

[19] Sanchez Fernandez I, Abend NS, Agadi S, An S, Arya R, Carpenter JL, et al. Gaps and opportunities in refractory status epilepticus research in children: a multi-center approach by the Pediatric Status Epilepticus Research Group (pSERG). Seizure. 2014;23(2):87-97.

[20] Stewart D. Accuracy of the Broselow tape for estimating paediatric weight in two Australian Emergency Departments: University of Sydney; 2009 [Accessed 27 September 2017, Master of Biostatistics]. Available from: https://ses.library.usyd.edu.au//bitstream/2123/6265/1/Declan%20Stewart%20WPP%20final.pdf

[21] Fein IA. Quality, patient safety, and culture: 'We have met the enemy and he is us'--Pogo (Walt Kelly, 1971). Critical Care Medicine. 2011;39(5):1196-7.

[22] Wells M, Goldstein L. Optimising emergency weight estimation in underweight and obese children: the accuracy of estimations of total body weight and ideal body weight. European Journal of Emergency Medicine. 2018; In Press.

[23] Young KD, Korotzer NC. Weight estimation methods in children: a systematic review. Annals of Emergency Medicine. 2016;68(4):441-451 e10.

[24] Wells M, Goldstein L, Bentley A. The accuracy of emergency weight estimation systems in children - a systematic review and meta-analysis. International Journal of Emergency Medicine. 2017;10(29):1-43. https://doi.org/10.1186/s12245-017-0156-5.

[25] Wells M, Goldstein L, Bentley A. A systematic review and meta-analysis of the accuracy of weight estimation systems used in paediatric emergency care in developing countries. African Journal of Emergency Medicine. 2017;7:S36-S54.

[26] Wells M, Barnes L, Vincent-Lambert C. Paediatric weight estimation practices of advanced life support providers in Johannesburg, South Africa. African Journal of Emergency Medicine. 2018;8(2):51-54.

[27] Huybrechts I, Himes JH, Ottevaere C, De Vriendt T, De Keyzer W, Cox B, et al. Validity of parent-reported weight and height of preschool children measured at home or estimated without home measurement: a validation study. BMC Pediatrics. 2011;11:63.

[28] Twarog JP, Politis MD, Woods EL, Daniel LM, Sonneville KR. Is obesity becoming the new normal? Age, gender and racial/ethnic differences in parental misperception of obesity as being 'About the Right Weight'. International Journal of Obesity (London). 2016;40(7):1051-1055.

[29] Robinson E. Overweight but unseen: a review of the underestimation of weight status and a visual normalization theory. Obesity Reviews. 2017;18(10):1200-1209.

[30] Clark MC, Lewis RJ, Fleischman RJ, Ogunniyi AA, Patel DS, Donaldson RI. Accuracy of the Broselow Tape in South Sudan, "The Hungriest Place on Earth". Academic Emergency Medicine. 2016;23(1):21-28.

[31] Batmanabane G, Kumar Jena P, Dikshit R, Abdel-Rahman SM. Using the Mercy Method for weight estimation in Indian children. Global Pediatric Health. 2015; January-December:1-7.

[32] A Manirafashaa, S Yi, GN Cattermole. Estimating children's weight in a Rwandan emergency centre. African Journal of Emergency Medicine. 2018; In Press.

[33] Samerchua A, Suraseranivongse S, Komoltri C. A comparison of pediatric weight estimation methods for emergency resuscitation. Pediatric Emergency Care. 2017;Apr 11. doi: 10.1097/PEC.0000000000001137. [Epub ahead of print].

[34] Geduld H, Hodkinson PW, Wallis LA. Validation of weight estimation by age and length based methods in the Western Cape, South Africa population. Emergency Medicine Journal. 2011;28(10):856-860.

[35] Akabarian S, Vahedparast H, Ravanipour M, Mirzaei K. Comparison of Broselow tape measurements versus mother estimations of pediatric weights. Iranian South Medical Journal. 2013;16(2):138-143.

[36] Nieman CT, Manacci CF, Super DM, Mancuso C, Fallon WF, Jr. Use of the Broselow tape may result in the underresuscitation of children. Academic Emergency Medicine. 2006;13(10):1011-9.

[37] Dicko A, Alhousseini ML, Sidibé B, Traoré M, Abdel-Rahman SM. Evaluation of the Mercy weight estimation method in Ouelessebougou, Mali. BMC Public Health. 2014;14:270.

[38] Kleinman ME, Chameides L, Schexnayder SM, Samson RA, Hazinski MF, Atkins DL, et al. Pediatric advanced life support: 2010 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Pediatrics. 2010;126(5):e1361-99.

[39] Wells M, Kramer E. Optimizing emergency drug dosing in children. Academic Emergency Medicine. 2008;15(12):1325.

[40] Wells M, Goldstein L, Bentley A. Accuracy of weight estimation by the Broselow tape is substantially improved by including a visual assessment of body habitus. Pediatric Research. 2018;83(1):83-92.

[41] Tanner D, Negaard A, Huang R, Evans N, Hennes H. A prospective evaluation of the accuracy of weight estimation using the Broselow Tape in overweight and obese pediatric patients in the Emergency Department. Pediatric Emergency Care. 2016;33(10):675-678.

[42] Advanced Life Support Group. Advanced Paediatric Life Support: The Practical Approach (APLS) 6th Edition. New York: BMJ Books - Publisher: John Wiley & Sons (Wiley-Blackwell); 2016.

[43] Marlow R, Lo D, Walton L. Accurate paediatric weight estimation by age: mission impossible? Archives of Disease in Childhood. 2011;96(Suppl 1):A1-A2.

[44] Cattermole GN, Graham CA, Rainer TH. Age-based weight prediction is imprecise. Emergency Medicine Journal. 2011;29(1):82-83.

[45] Erker CG, Santamaria M, Moellmann M. Size does matter--age-related weight estimation in "tall n' thin" and "tiny n' thick" children and a new habitus-adapted alternative to the EPLS-formula. Resuscitation. 2014;85(9):1174-1178.

[46] Appelbaum N, Clarke J, Maconochie I, Darzi A. Paediatric weight estimation by age in the digital era: optimising a necessary evil. Resuscitation. 2018;122:29-35.

[47] Carley S. APLS weight estimation – don't do it (well almost never). 2013 [Accessed March 2018]. Available from: http://stemlynsblog.org/apls-estimation-formulas-do-not-safely-predict-weight-in-uk-children-st-emlyns/#comments

[48] Marikar D, Varshneya K, Wahid A, Apakama O. Just too many things to remember? A survey of paediatric trainees' recall of Advanced Paediatric Life Support (APLS) weight estimation formulae. Archives of Disease in Childhood. 2013;98(11):921.

[49] Britnell S, Koziol-McLain J. Weight estimation in paediatric resuscitation: a hefty issue in New Zealand. Emergency Medicine Australasia. 2015;27(3):251-256.

[50] Downing A, Rudge G. A study of childhood attendance at emergency departments in the West Midlands region. Emerg Med J. 2006;23(5):391-3.

[51] Cattermole GN, Leung PY, Graham CA, Rainer TH. Too tall for the tape: the weight of schoolchildren who do not fit the Broselow tape. Emergency Medicine Journal. 2014;31(7):541-544.

[52] Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 1999-2016.: National

Health and Nutrition Examination Survey Data; 2016. Available from: https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx

[53] Wells M, Goldstein L, Bentley A. A validation study of the PAWPER XL tape: accurate estimation of both total and ideal body weight in children up to 16 years of age. Trauma and Emergency Care. 2017;2(4):1-8.

[54] Pinchevsky LE, Pesaturo KA, Smith BS, Hartman CA. Pilot comparison of three cardiopulmonary resuscitation medication dosing strategies in overweight children. Journal of Pediatric Pharmacology and Therapeutics. 2010;15(4):282-9.

[55] Global Burden of Disease Study 2011. Global Burden of Disease Study 2013 (GBD 2013) Obesity Prevalence 1990-2011. Seattle, United States: Institute for Health Metrics and Evaluation (IHME); 2014.

[56] Gardner RM, Brown D. Body image assessment: A review of figural drawing scales. Personality and Individual Differences. 2010;48:107-111.

[57] Garland JS, Kishaba RG, Nelson DB, Losek JD, Sobocinski KA. A rapid and accurate method of estimating body weight. American Journal of Emergency Medicine. 1986;4(5):390-393.

[58] Schmidt AR, Buehler PK, Meyer J, Weiss M, Schmitz A, Both CP. Length-based body weight estimation in paediatric patients: the impact of habitus-a clinical observational trial. Acta Anaesthesiologica Scandinavica. 2018.

[59] Schmidt AR, Haas T, Buehler KP, Ruetzler K, Weiss M, A S. A comparison of Broselow tape and pediatric emergency ruler. Accuracy of weight estimation. Notfall Rettungsmedizin. 2015;19:129-135.

[60] Yamamoto LG, Inaba AS, Young LL, Anderson KM. Improving length-based weight estimates by adding a body habitus (obesity) icon. American Journal of Emergency Medicine. 2009;27(7):810-815.

[61] Leykin Y, Miotto L, Pellis T. Pharmacokinetic considerations in the obese. Best Practice & Research Clinical Anaesthesiology. 2011;25(1):27-36.

[62] Harskamp-van Ginkel MW, Hill KD, Becker K, Testoni D, Cohen-Wolkowiez M, Gonzalez D, et al. Drug dosing and pharmacokinetics in children with obesity: a systematic review. Journal of the American Medical Association Pediatrics. 2015;169(7):678-85.

[63] Kendrick JG, Carr RR, Ensom MH. Pediatric obesity: pharmacokinetics and implications for drug dosing. Clinical Therapeutics. 2015;37(9):1897-923.

[64] Ross EL, Heizer J, Mixon MA, Jorgensen J, Valdez CA, Czaja AS, et al. Development of recommendations for dosing of commonly prescribed medications in critically ill obese children. American Journal of Health-System Pharmacy. 2015;72(7):542-556.

[65] Rowe S, Siegel D, Benjamin DK, Jr., Best Pharmaceuticals for Children Act - Pediatric Trials Network Administrative Core C. Gaps in drug dosing for obese children: a systematic review of commonly prescribed emergency care medications. Clinical Therapeutics. 2015;37(9):1924-1932.

[66] Dieckmann RA. The dilemma of paediatric drug dosing and equipment sizing in the era of patient safety. Emergency Medicine Australasia. 2007;19:490-493.

[67] Gardner RM, Stark K, Jackson NA, Friedman BN. Development and validation of two new scales for assessment of body-image. Perceptual and Motor Skills. 1999;89(3 Pt 1):981-93.

[68] Gardner RM, Jappe LM, Gardner L. Development and validation of a new figural drawing scale for body-image assessment: the BIAS-BD. Journal of Clinical Psychology. 2009;65(1):113-22.

[69] Yepes M, Viswanathan B, Bovet P, Maurer J. Validity of silhouette showcards as a measure of body size and obesity in a population in the African region: A practical research tool for general-purpose surveys. Population Health Metrics. 2015;13:35.

[70] Truby H, Paxton SJ. Development of the Children's Body Image Scale. British Journal of Clinical Psychology. 2002;41(Pt 2):185-203.

[71] Truby H, Paxton SJ. The Children's Body Image Scale: reliability and use with international standards for body mass index. British Journal of Clinical Psychology. 2008;47(Pt 1):119-24.

[72] Britnell S, Taylor S, Koziol-McLain J. Emergency weight estimation lookup tables for New Zealand children aged 5-10 years. Emergency Medicine Australasia. 2016;28(5):558-63.

[73] Clasey JL, Bradley KD, Bradley JW, Long DE, Griffith JR. A new BIA equation estimating the body composition of young children. Obesity (Silver Spring). 2011;19(9):1813-7.

[74] Abdel-Rahman SM, Jacobsen R, Watts JL, Doyle SL, O'Malley DM, Hefner TD, et al. Comparative performance of pediatric weight estimation techniques:

a human factor errors analysis. Pediatric Emergency Care. 2017;33(8):548-552.

[75] Cornelissen KK, Gledhill LJ, Cornelissen PL, Tovee MJ. Visual biases in judging body weight. British Journal of Health Psychology. 2016;21(3):555-69.

[76] Cattermole GN, Leung PY, Mak PS, Graham CA, Rainer TH. Mid-arm circumference can be used to estimate children's weights. Resuscitation. 2010;81(9):1105-10.

[77] Whitfield KC, Wozniak R, Pradinuk M, Karakochuk CD, Anabwani G, Daly Z, et al. Anthropometric measures are simple and accurate paediatric weight-prediction proxies in resource-poor settings with a high HIV prevalence. Archives of Disease in Childhood. 2017;102(1):10-16.

[78] Wells M, Goldstein L, Bentley A. Development and validation of a method to estimate bodyweight in critically ill children using length and mid-arm circumference measurements - the PAWPER XL-MAC system. South African Medical Journal. 2017;107(11):1015-1021.

[79] Georgoulas V, Wells M. The PAWPER tape and the Mercy Method outperform other methods of weight estimation in children in South Africa. South African Medical Journal. 2016;106(9):933-939.

[80] WHO Multicentre Growth Reference Study Group. Reliability of anthropometric measurements in the WHO Multicentre Growth Reference Study. Acta Paediatrica. 2006;450(Suppl):38-46.

[81] Hohenhaus SM, Frush KS. Pediatric patient safety: common problems in the use of resuscitative aids for simplifying pediatric emergency care. Journal of Emergency Nursing. 2004;30(1):49-51.

[82] Luten R. Error and time delay in pediatric trauma resuscitation: Addressing the problem with color-coded resuscitation aids. Surgical Clinics of North America. 2002;82(2):303-314.

[83] Marlow RD, Wood DLB, Lyttle MD. Comparing the usability of paediatric weight estimation methods: a simulation study. Archives of Disease in Childhood. 2018.

[84] Hohenhaus SM, Frush KS. Pediatric patient safety in resuscitation simulations: lessons learned from studying clinician behavior. Pediatric Emergency Care. 2005;21(1):71-75.

[85] Hunt EA, Hohenhaus SM, Luo X, Frush KS. Simulation of pediatric trauma stabilization in 35 North Carolina emergency departments: identification of targets for performance improvement. Pediatrics. 2006;117(3):641-648.

[86] Gaca AM, Frush DP, Hohenhaus SM, Luo X, Ancarana A, Pickles A, et al. Enhancing pediatric safety: using simulation to assess radiology resident preparedness for anaphylaxis from intravenous contrast media. Radiology. 2007;245(1):236-44.

[87] Lerner C, Gaca AM, Frush DP, Hohenhaus S, Ancarana A, Seelinger TA, et al. Enhancing pediatric safety: assessing and improving resident competency in life-threatening events with a computer-based interactive resuscitation tool. Pediatric Radiology. 2009;39(7):703-9.

[88] Walsh KE, Adams WG, Bauchner H, Vinci RJ, Chessare JB, Cooper MR, et al. Medication errors related to computerized order entry for children. Pediatrics. 2006;118(5):1872-1879.

[89] Yamamoto L, Kanemori J. Comparing errors in ED computer-assisted vs conventional pediatric drug dosing and administration. American Journal of Emergency Medicine. 2010;28(5):588-92.

[90] Abdel-Rahman SM, Ahlers N, Holmes A, Wright K, Harris A, Weigel J, et al. Validation of an improved pediatric weight estimation strategy. Journal of Pediatric Pharmacology and Therapeutics. 2013;18(2):112-121.

[91] Arnholt AM, Duval-Arnould JM, McNamara LM, Rosen MA, Singh K, Hunt EA. Comparatively evaluating medication preparation sequences for treatment of hyperkalemia in pediatric cardiac arrest: a prospective, randomized, simulation-based study. Pediatric Critical Care Medicine. 2015;16(7):e224-30.

[92] Abdel-Rahman SM, Paul IM, James LP, Lewandowski A. Evaluation of the Mercy TAPE: performance against the standard for pediatric weight estimation. Annals of Emergency Medicine. 2013;62(4):332-339.e6.

[93] Wells M, Goldstein L, Bentley A. The accuracy of paediatric weight estimation during simulated emergencies: the effects of patient position, patient cooperation and human errors. African Journal of Emergency Medicine. 2017; In Press.

[94] TSG Associates. Smart Tape providing evidence based triage for injured children 2018 [Accessed 30 July 2018]. Available from: http://www.smartmci.com/products/triage/smart_tape.php

[95] Both CP, Schmitz A, Buehler PK, Weiss M, Schmidt AR. How accurate are pediatric emergency tapes? A comparison of 4 emergency tapes with different length-based weight categorization. Pediatric Emergency Care. 2017.

[96] Lowe CG, Campwala RT, Ziv N, Wang VJ. The Broselow and Handtevy Resuscitation Tapes: a comparison of the performance of pediatric weight prediction. Prehospital and Disaster Medicine. 2016;31(4):364-75.

[97] Wells M, Goldstein L, Bentley A, Basnett S, Monteith I. The accuracy of the Broselow tape as a weight estimation tool and a drug-dosing guide - a systematic review and meta-analysis. Resuscitation. 2017;121:9-17. http://dx.doi.org/10.1016/j.resuscitation.2017.09.026.

[98] Luten R, Wears RL, Broselow J, Croskerry P, Joseph MM, Frush K. Managing the unique size-related issues of pediatric resuscitation: reducing cognitive load with resuscitation aids. Academic Emergency Medicine. 2002;9(8):840-7.

[99] Slishman S, Sapien R, Crandall CS. Introducing a simple, weight-based, color-coded, medication dosing device. Pediatric Emergency Care. 2002;18(3):212-215.

[100] Moreira ME, Hernandez C, Stevens AD, Jones S, Sande M, Blumen JR, et al. Color-coded prefilled medication syringes decrease time to delivery and dosing error in simulated Emergency Department pediatric resuscitations. Annals of Emergency Medicine. 2015;66(2):97-106 e3.

[101] Stevens AD, Hernandez C, Jones S, Moreira ME, Blumen JR, Hopkins E, et al. Color-coded prefilled medication syringes decrease time to delivery and dosing errors in simulated prehospital pediatric resuscitations: a randomized crossover trial. Resuscitation. 2015;96:85-91.

[102] Broselow J. From humble beginnings: the birth of the Broselow Tape 2012 [Accessed March 2018]. Available from: http://epmonthly.com/article/from-humble-beginnings-the-birth-of-the-broselow-tape/

[103] Luten R, Broselow J. Medication guide for pediatric resuscitation and emergency infusions. USA: Armstrong Medical Industries Inc; 2017.

[104] Fineberg SL, Arendts G. Comparison of two methods of pediatric resuscitation and critical care management. Annals of Emergency Medicine. 2008;52(1):35-40 e11.

[105] Shah AN, Frush K, Luo X, Wears RL. Effect of an intervention standardization system on pediatric dosing and equipment size determination: a crossover trial

involving simulated resuscitation events. Archives of Pediatric and Adolescent Medicine. 2003;157(3):229-36.

[106] Wells M, Goldstein L, Bentley A. The use of body habitus reference images improves the ability of novices to accurately estimate children's weight using the PAWPER XL tape system. The Journal of Emergency Medicine. 2017;54(2):165-175.

[107] Wells M, Goldstein L, Bentley A. The accuracy of paediatric weight estimation during simulated emergencies: the effects of patient position, patient cooperation and human errors. African Journal of Emergency Medicine. 2018;8(2):43-50.

[108] Frush KS. Medication errors in pediatric emergency care: developing a national standard? Annals of Emergency Medicine. 2007;50(4):369-370.

[109] Morgan N, Luo X, Fortner C, Frush K. Opportunities for performance improvement in relation to medication administration during paediatric stabilisation. Quality & Safety in Health Care. 2006;15:179-183.

[110] McDaniel-Hohenhaus S, Frush K. Development of an e-learning program for use of the Broselow-Luten pediatric system. Journal of Emergency Nursing. 2003;29(5):417.

[111] Mayes J, White A. How smartphone technology is changing healthcare in developing countries. The Journal of Global Health. 2016;7(1):36-38.

Chapter 12

12.1 Integrated conclusions

The aim of this research was to systematically evaluate the accuracy of existing methods of weight estimation, to refine the functioning of the PAWPER system and validate its enhancements, as well as to evaluate the usability and resilience of various weight estimation systems, including a mobile phone application.

In terms of the accuracy of the weight estimation studies evaluated in this thesis, it was clear that the dual length- and habitus-based systems were the most accurate, along with parental estimates of weight. The Mercy method and the PAWPER tape maintained their accuracy across different populations and in most subgroups of body composition. These two methods could truly be regarded as global weight estimation systems. However, further research is required to establish whether parental estimates of weight are accurate in all populations, especially those from low-income countries. It was also clear that the accuracy of the one-dimensional age- or length-based systems were substantially inferior to that of the Mercy method and the PAWPER tape. The age- and length-based systems showed significant disparities in their performance in different populations. They should not be used in clinical practice unless no other method is available.

The evidence presented in this thesis showed that the pre-calculated drug dose information printed on the Broselow tape did not contain the degree of detail needed for drug dose calculation, preparation and administration. The necessity to have additional reference material to supplement the Broselow tape negated the value of the pre-calculated doses on the tape. The poor accuracy of the tape, its inferior usability in uncooperative children and non-supine children and the questionable value of the information printed on the tape limit the value of the tape significantly. International guidelines which recommend the use of length-based tapes with pre-calculated doses in paediatric emergencies should be reconsidered.

The studies included in this thesis also established that the newly-developed PAWPER XL tape estimated weight more accurately than the original PAWPER tape. In addition, the PAWPER XL tape provides accurate weight estimations in taller children, up to 180 cm in length, and it was significantly more accurate than the original PAWPER tape in obese and severely obese children. The PAWPER XL tape was also able to produce accurate estimations of IBW in obese children.

The two studies that evaluated the more objective methods of assessing habitus had encouraging outcomes. The figural reference image method for assisting habitus assessment was significantly more accurate than the gestalt method described for use with the original PAWPER tape. The system was also more popular with the users. This system has the advantage of standardising visual habitus assessment in different settings. The most objective weight estimation method evaluated in the thesis was the PAWPER XL-MAC system, which used mid-arm circumference measurements to define habitus categories. The PAWPER XL-MAC system showed a high level of accuracy of weight estimation in datasets from populations with differing prevalences of underweight and obese children. The accuracy was similar to results obtained with the PAWPER XL tape in the same datasets. This system also has the advantage of near-complete objective and standardised assessment of habitus but will need to be evaluated under real-life conditions.

Testing the usability of four different weight estimation systems under adverse conditions revealed that both patient factors and human factors significantly compromised all aspects of the performance of the Broselow tape. The PAWPER XL tape was more resilient to unfavourable patient factors than the Mercy method and was also significantly less vulnerable to human factor errors. The findings suggested that the PAWPER XL tape will retain its accuracy during real life emergency care. The findings were also strongly suggestive of the need for appropriate training for individuals using weight estimation systems during emergencies.

The use of the mobile phone application was moderately successful: the weight estimation produced by the app was accurate, but the use of new technology produced unexpected challenges which affected its usability. There is enormous

potential for mobile phone applications to improve the accuracy of weight estimation and drug dosing while reducing the cognitive load that would be experienced without the assistance of technology

The work from this thesis has contributed in a small way to the understanding of emergency weight estimation in children. There was a dearth of high quality evidence to be found on many aspects of emergency weight estimation and a great deal of further research is still required.

12.2 Recommendations and further research

12.2.1 Recommendations

The following recommendations, based on the research outcomes presented in this thesis, are intended to promote evidence-based paediatric emergency care. If adopted, these guidelines would allow healthcare providers to provide the best possible care for critically ill or injured children in need of urgent treatment.

The classification of recommendations used in this chapter is modified from the guidelines produced by the American Heart Association (shown in Table 12-1) [1].

Table 12-1 The classification of recommendations and level of evidence used in this chapter.

Classification of Recommendations	
Class I	Conditions for which there is evidence, general agreement, or both that a given procedure or treatment is useful and effective.
Class II	Conditions for which there is conflicting evidence, a divergence of opinion, or both about the usefulness or efficacy of a procedure or treatment.
Class IIa	Weight of evidence or opinion is in favour of usefulness or efficacy.
Class IIb	Usefulness or efficacy is less well established by evidence or opinion.
Class III	Conditions for which there is evidence, general agreement, or both that the procedure or treatment is not useful or effective and, in some cases, may be harmful.
Level of Evidence (LOE)	
LOE A	Data derived from multiple randomised clinical trials
LOE B	Data derived from a single randomised trial or nonrandomised studies
LOE C	Consensus opinion of experts

1. In scenarios where there is some time to acquire a weight estimation, but a child cannot be weighed, an estimation of weight from the Mercy method, the PAWPER XL tape or a parent should be used. **Class I LOE B**
2. In scenarios where there is very limited time to acquire a weight estimation, an estimation of weight from the PAWPER XL tape or a parent may be used. **Class IIb LOE C**
3. The PAWPER XL tape should be used together with figural reference images to assess habitus in the early stages of the resuscitation. **Class IIb LOE C**
4. The PAWPER XL-MAC system could be appropriate for weight estimation during emergency care. **Class IIb LOE C**
5. In general, weight estimations obtained using age-based formulas and length-based tapes should not be used, except in select populations or subgroups in which they have been shown to be accurate. **Class III LOE B**
6. All healthcare providers who provide emergency care to children, should be well-trained and well-practised with whichever weight estimation system they elect to use. **Class I LOE B**

7. Complete resuscitation aids must be used during paediatric emergency care. They must integrate with a weight estimation system, must contain information on drug doses as well as instructions for preparation and administration and must eliminate all calculations. **Class I LOE B**
8. The Broselow tape is not a complete resuscitation aid and should not be used for this purpose. **Class III LOE C**
9. Technological aids to weight estimation and drug dosing, especially mobile phone applications, should be developed for emergency use. **Class IIb LOE C**

12.2.2 Future research

Despite the major advances in knowledge in the last five years, especially in terms of the development of the Mercy method and the PAWPER tape systems, much research is still needed. This research should focus on improving the accuracy of weight estimation in underweight and obese children, on increasing the usability of weight estimation systems and on integrating weight estimation and drug dosing aids.

Perhaps most importantly, the use of mobile phone technology needs to be explored expeditiously and its use expanded, training in weight estimation needs to become commonplace and the use of accurate weight estimation methods needs to be accelerated.

Some of the noteworthy fields for future research, arising from ideas generated in this thesis, are shown in Table 12-2.

Table 12-2 Potential future research questions.

Future research possibilities
Accuracy of weight estimation systems
Are virtual weight estimation studies valid?
Are parental estimates of weight in rural and limited resource communities valid?
Does the accuracy of weight estimation change outcome: is an accurate estimation of weight really necessary?
How accurate should weight estimation systems be?

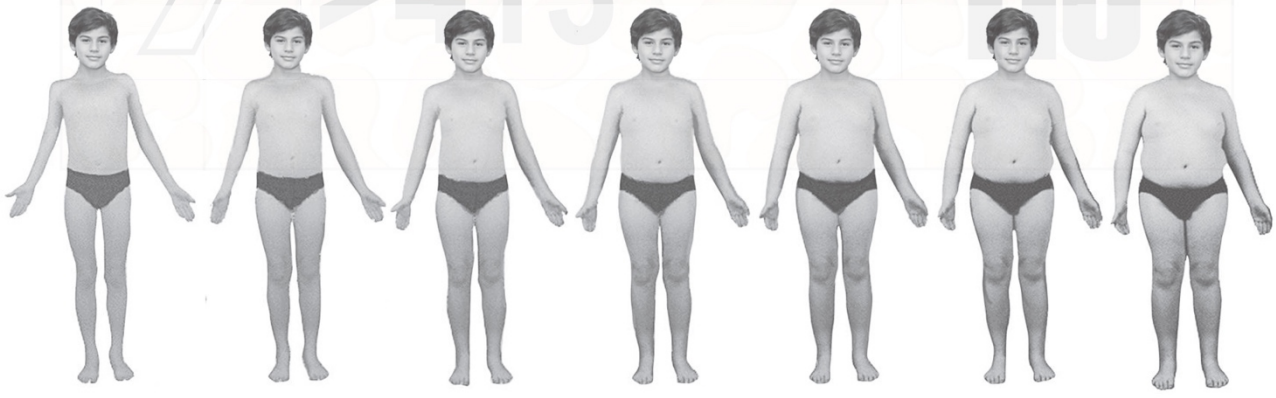
What affects the accuracy of visual estimation of habitus?
How many habitus score categories are required for accurate weight estimation?
Can technology such as point-of care ultrasound be used to aid weight estimation?
Can technology such as automated length measurement and habitus assessment be used to aid weight estimation?
How can accessibility to the PAWPER systems be increased to healthcare providers across the world, without becoming cost-prohibitive?
Would the PAWPER XL-MAC system be accurate in low- and middle-income countries?
Usability of weight estimation systems
Does confidence in weight estimation method improve the initial delivery of emergency care?
How does weight estimation and drug dosing impact on cognitive stress and load during emergencies?
Which weight estimation systems are actually used in the prehospital environment and in the emergency departments across the world?
Does the physical size of a weight estimation tape influence the accuracy of weight estimation?
Would a staged, or step-by-step, weight estimation algorithm be useful?
What training is best for weight estimation and drug dosing?
How rapidly does skill erosion take place for weight estimation methods?
Would there be any benefit to incorporating weight estimation training into advanced life support training course paediatric simulations.
Integration of weight estimation and drug dosing
What dose error should be considered a dose error?
What is known about paediatric pharmacology and drug dosing in emergencies?
Can all drugs be titrated to effect, eliminating the need for accurate weight estimation?
Could the Mercy method be adapted to provide an estimation of IBW in obese children?
Which is preferable as a dose scalar for obese children: IBW or LBW?
Can technology such as mobile phone apps be used to integrate weight estimation and drug dosing effectively?
How do compounded or cumulative errors affect the accuracy of drug dosing?

12.3 References

[1] Budoff MJ, Achenbach S, Blumenthal RS, Carr JJ, Goldin JG, Greenland P, et al. Assessment of coronary artery disease by cardiac computed tomography: a scientific statement from the American Heart Association Committee on Cardiovascular Imaging and Intervention, Council on Cardiovascular Radiology and Intervention, and Committee on Cardiac Imaging, Council on Clinical Cardiology. Circulation. 2006;114(16):1761-91.

1	<257	55
2	257-282	63
3	283-326	70
4	327-360	87
5	361-395	95
6	396-413	105
7	>413	116

Appendixes



Appendix 1

Approved PhD protocol

The protocol that was approved by the Postgraduate Committee of the Faculty of Health Sciences of the University of the Witwatersrand is included in this appendix.

The focus of the research evolved slightly with the availability of new information in the international literature. The content of the thesis was also limited due to the magnitude of the data produced. During the research, the lack of consensus on the accuracy of weight estimation systems became apparent. The systematic reviews were therefore conducted and included in the thesis. During the research process, information on the accuracy of the Mercy method was published. Given the good performance of the Mercy method and its standing as one of the best weight estimation systems in the world, it was included in much of this research as well.

Specific notes:

1. Data on fat mass and fat free mass were collected using DXA but this aspect was not included in the thesis but will be addressed in later publications.
2. The emphasis on children with dehydration and over-hydration was removed both because of very low numbers of children presenting with these conditions and a shift of focus in the research.
3. Data were collected using bioelectrical impedance and point-of-care ultrasound, but this was not included in the thesis. It has been published.
4. The algorithm mentioned in the protocol was created but was not included in the thesis. A modification of this algorithm has been published.
5. Since clarity had emerged from other research on the ideal number of habitus score categories, this was not pursued in this thesis.
6. Although data was collected on the ability of healthcare providers to guess BMI, this was not included in the thesis.

ESTIMATING CHILDREN'S WEIGHT IN THE EMERGENCY DEPARTMENT USING A LENGTH-BASED, HABITUS-MODIFIED SYSTEM TO INCREASE THE ACCURACY OF TOTAL BODY WEIGHT ESTIMATION: THE UTILITY OF THE PAWPER TAPE IN PAEDIATRIC EMERGENCY MEDICINE

A research proposal presented to the

Division of Emergency Medicine,

Faculty of Health Sciences,

University of the Witwatersrand

For fulfilment of the requirements for the degree

Doctor of Philosophy

by

Dr M Wells

8523767

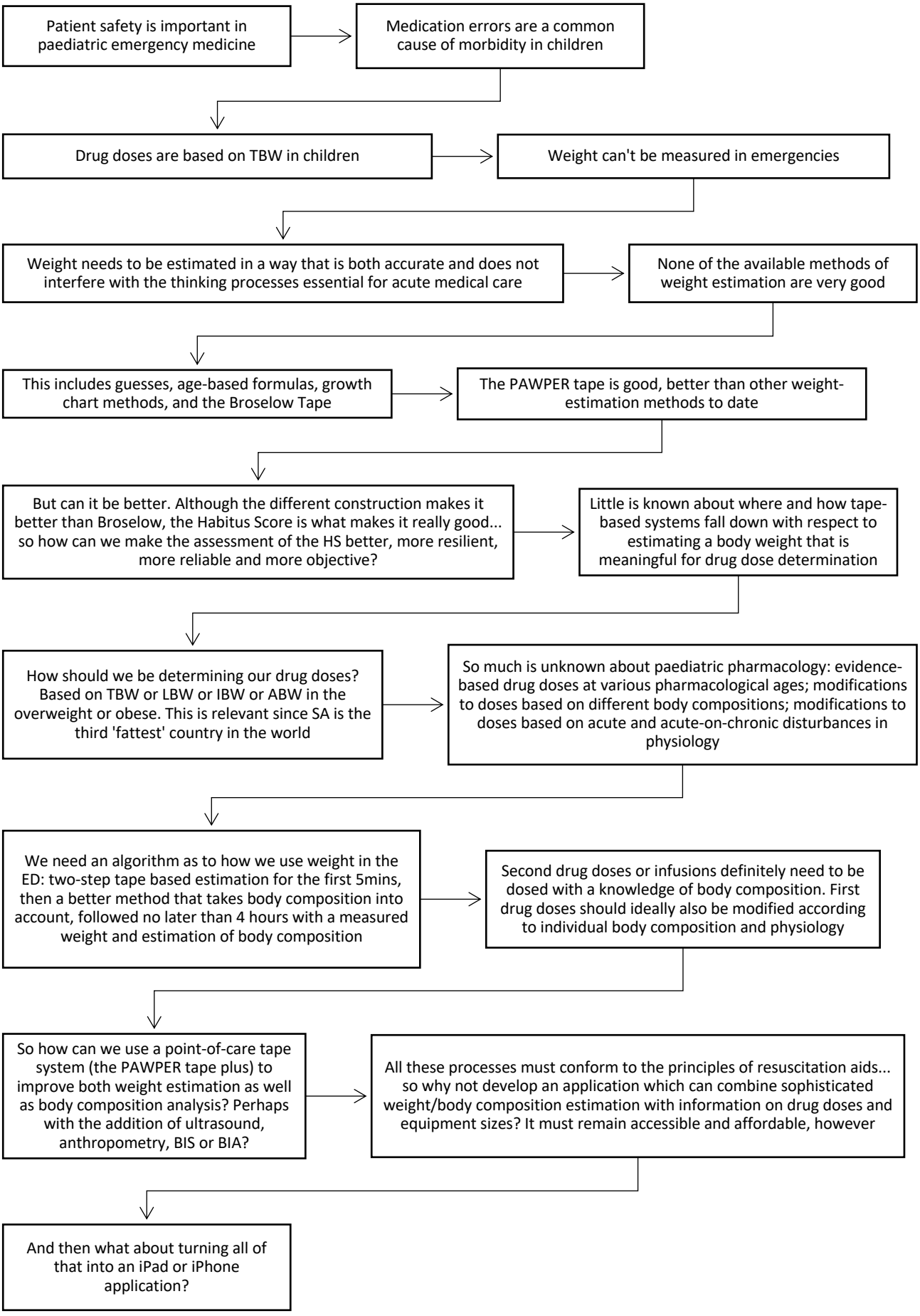
Supervisor

Dr A Bentley

_____

INTRODUCTION

Outline: the following topics will be covered briefly in order to put the research into perspective:



Patient safety – a paramount concern

The significance of medical errors has come under the spotlight in recent times with the Institute of Medicine (IOM) in the United States estimating in 2001 that perhaps nearly 100,000 people die each year as a result of medical errors, of which the majority are medication errors.[1] The IOM published further comments in 2006 that medical errors in children were more common than in adults and that a system should be developed to reduce preventable medication errors.[2] This matter generated such attention that the Joint Commission for Accreditation of Healthcare Organisations (JCAHO) in the USA released a sentinel alert in April 2008 entitled "Preventing pediatric medication errors" with recommendations and guidelines for reducing medication errors in children (see Appendix) and in 2009 the National Health Service in the UK produced a document entitled "Review of patient safety for children and young people" which highlighted some of the patient safety issues for children that had been identified. Second only to failures in recognising the severity of illness in children, was the high rate of reported medication administration errors, particularly dosing errors.

It is the now the middle of the second decade since the medical world was galvanised into discussions about its "discovery of the epidemic of medical mistakes" and many experts have declared that the lack of progress has been "astounding", or even that we have experienced "the end of the quality improvement movement". Two editorials in 2011 even went so far as to headline their articles with Walt Kelly's famous 1971 cartoon quip, "We have met the enemy, and he is us".[3, 4] Clearly much has still to change, not least doctors' and healthcare organisations' attitudes towards patient safety.

In most hospitals, perhaps understandably, the paediatric and emergency departments generate the most medication errors.[5] Accurate drug dose calculation and delivery in children and infants can be challenging at the best of times in the ED, but in the stressful circumstances of a paediatric emergency or resuscitation, however, it is even more difficult and susceptible to error.[6, 7] There is a growing concern that, in children, mistakes in the process of medication prescription and administration are more common than in adults, and are more significant, or have a higher potential for harm, than was previously realised.[8] It is for this reason that it is essential for any doctor who is involved in the emergency management of children to have a knowledge of the nature of medication errors that occur in the ED, in order to develop a strategy to eliminate (or at least minimise) potential resultant harm to the patient. It requires a system designed to reduce medication-related harm to create a "culture of safety" for infants and children in the ED.[9] The American Academy of Pediatrics updated a decade-old policy statement in 2011 entitled "Principles of pediatric patient safety: reducing harm due to medical care" in which they emphasised the importance of the establishment of a culture of safety, leadership, the use of information technology, the generation of patient-safety goals, the use decision-support systems and the elimination of medication errors (and especially dosing errors).[10]

What errors in the administration of medication to children can occur in the ED?

Case scenario

> *A 5 year old boy with severe haemophilia was knocked off his bicycle by a car. He is brought into the ED by ambulance with a Glasgow Coma Score of 7 out of 15, a systolic blood pressure of 75mmHg, a heart rate of 160 beats per minute and a fractured right femur.*

While it might seem simple to “give etomidate” as part of a rapid sequence intubation for initial airway management in this scenario, the potentials for errors standing between this child and the successful administration of medications (and other interventions) that might prove life-saving are immense. The management team confronted by this case must:

- Identify the initial management priorities.
- Identify the appropriate drug to administer.
- Determine the correct dose in milligram per kilogram format, including modifications to the dose relating to age, critical illness, obesity, other medications and possible drug interactions.
- Obtain a measure of the child’s mass by measurement or estimation.
- Correctly calculate the total dose to administer from the base milligrams per kilogram dosage and the child’s mass.
- Communicate the prescription verbally, in writing or via computer to the nurse or pharmacist who will prepare the medication.
- The nurse (or pharmacist) must accurately determine and perform the required dilution of the drug.
- The nurse (or pharmacist) must correctly calculate the volume of drug to administer.
- The nurse or doctor must administer the drug correctly.

Each of these steps must be correctly completed to ensure that no medication error occurs. Furthermore, this process must be repeated for additional necessary medications, each time avoiding the pitfalls of potential errors.

Doctors and nursing staff members who are not experienced in paediatric emergency medicine or the management of paediatric emergencies are especially at risk of making medication-related errors.[8] Studies have shown that 10% to 15% of paediatric medication orders are erroneous, and that nearly one third of medication errors lead to harm in children.[11] The potential for harm is higher in smaller children and infants and in those critically ill or injured, possibly because of the administration of multiple medications, limited time to double-check doses and altered pharmacokinetics and pharmacodynamics resulting from the physiological insults.[1]

Drug dose recommendations in children are generally based on body weight (or very occasionally body surface area) and therefore appropriate drug administration requires an accurate measured or estimated body weight and a dose calculation, dilution and delivery that is free of errors. Not surprisingly, the most common medication errors in children in the ED are dosing errors, which can originate at any stage of the process – calculation, prescription, preparation or administration.[12]

The emergency physician must first obtain the correct milligram per kilogram dose either from memory (which may be very unreliable) or from appropriate reference material, and then determine if any age or physiology-related adjustments need to be made.

This base dose then needs to be converted to an actual total dose calculated from the patient's body weight. Ideally body weight should be measured (the child weighed on a scale) for drug dose calculations. In practice it is often impossible to measure weight in children who are critically ill or injured and undergoing resuscitative treatment and if weight cannot be measured then it must be estimated since all drug dose calculations make use of body weight as an index (see below).

Once the weight has been determined the total dose to be administered needs to be calculated. This simple arithmetic exercise has been shown to be very poorly performed even by experienced doctors and order of magnitude (tenfold) errors in drug doses calculations are common in children.

After the dose has been determined, this must be communicated to the staff member who will prepare the drug for administration. This may occur via a verbal instruction, handwritten prescription or computerised system. Misunderstandings are frequent following verbal instructions. Illegible handwriting, incorrect positioning of the decimal point, transcription errors and mental lapses (such as using milligrams instead of micrograms inappropriately) are common causes of errors.[13] Typographical errors mean that computerised systems are not immune from error either.

Once the drug order has been given, the drug must be correctly prepared, diluted and administered by the nurse or doctor. This process also requires the calculation of the millilitre volume of drug dilution to administer and is susceptible to error, especially in inexperienced hands. The wrong amount of drug, the wrong type or volume of diluent and the wrong volume administered to the patient are additional causes of errors.

In order to provide optimum medical care these processes must be swiftly and successfully completed at the same time as managing the intellectual and emotional stress of a difficult paediatric resuscitation.[11] Wrong drugs, wrong doses, wrong administration as well as delays in treatment caused by ponderous calculations may negatively impact management and ignorance of drug doses or the fear of errors may lead to the omission of necessary medications. [6, 14] Critical thinking time may well be wasted on the calculation of drug dosages and the determination of equipment sizes rather than focused on the medical management of the patient. Systems need to be in place to assist the paediatric emergency team in the process of TBW estimation, drug dose calculation (including necessary modifications for body composition and acute physiological state) and drug delivery.

How can we estimate weight in the ED?

A number of different methods have been used to estimate the weight of children and infants in these situations when drugs need to be administered and weight cannot be measured. Some of the methods that have been used include:

- Estimation by the treating doctor (a guess at the weight);
- Estimation by the parent/s (a guess guided by recent weight measurements, if any);
- Prediction using age-based formulas e.g. weight = 8 + (age x 2) or weight = 7 + (age x 3); with weight in kg and age in years;

- Growth chart methods – weight-for-age or weight-for-length
- Length-based prediction methods
    - Tape methods (eg Broselow Tape);
    - Tape-and-modified table methods eg devised weight estimation method (DWEM);
    - Tape-and-table methods (e.g. Kloeck Tape + American Heart Association (AHA) Emergency Cardiovascular Care (ECC) Booklet).

While none of these systems are perfect, some are considerably better than others. Ideally the objective in weight estimation is a system that reliably predicts weight within 10% of measured weight, across a wide spectrum of age groups and body habitus. Estimates of the patient's weight by healthcare workers are notoriously inaccurate, but estimates of the child's weight by the parents are often very accurate, and acceptable to be used in an emergency situation. Alternative methods if parents do not accompany the child include age-based formulas or length-based prediction methods such as the derived weight estimation method or the Broselow tape. Age based formulas vary greatly between different populations and are unable to account for variations between children who are large, average or small for their age.

The most widely used system currently available is the length-based tape system, most notably the "Broselow Pediatric Emergency Tape". This was the first device specifically designed for weight prediction in children in emergency situations, and has achieved world-wide popularity. Despite this, however, there are some problems with the Broselow tape system

- It is designed to underestimate (or more accurately, not to overestimate) weight. Therefore, unsurprisingly, it has been shown to underestimate weight, especially in obese children (an increasingly common problem in many countries worldwide, including South Africa).
- It has been shown to be less accurate in populations other than in the USA (eg Maori and Pacific Island populations).
- It contains drug information for drugs that are not available in South Africa.
- It contains drug information with doses in mass format (mg or µg), without reference to dilution of the drug.
- It is a multi-use device that is not readily sterilisable and therefore a potential fomite.
- It is expensive in South Africa (R500+).

In order to overcome some of these limitations (especially with under- or overestimating weight) Jeffrey Garland and collaborators[15] developed the devised weight estimation system (DWEM) in 1986. They created a table with different weights for each height cohort (5cm intervals) based on frame size: small frame, medium frame or large frame. This showed good predictive ability, but has not been used much in practice or research since then. Garland *et al* summed up the entire matter, however, in a letter entitled "Body habitus + height = accurate weight estimate": a concept which is even more true today with the PAWPER tape[16] than it was in 1989.[17] More recently Yamamoto et al[18] derived regression equations to estimate weight for different habitus categories (icons) which, although complex, showed an improvement on traditional methods. Although the PAWPER tape in its current form is accurate, it might be possible to

improve on this accuracy with modifications to the tape methodology, including the methods of assessing the habitus score.

What other methodology could add to the accuracy of weight-estimation by the PAWPER tape?

A two-stage weight estimation using length and body habitus might possibly be improved upon by the addition of another step using a variable that has been shown to be associated with body weight or body composition. Some of the variables that have been considered include:

- Anthropometry: leg length or relative leg length, foot length, mid-arm circumference, head circumference;
- Body composition assessment: skinfold thickness, BIA or BIS, ultrasound (A-mode or B-mode)

While many of these modalities have not been used in the setting of weight or body composition assessment for the purposes of drug dose determination, they may well be amenable for adaptation into a weight-estimation system. Every existing weight-estimation system today is used to estimate TBW, but not every drug should be dosed according to TBW in every patient. Dose modification according to body composition (primarily obesity) and the child's acute physiological state are also important factors to be considered.

How should we be determining our drug doses?

Not every drug dose should be calculated according to TBW in children. Since every drug has different pharmacology and every patient is different, drug doses should be individualised according to each drug to be used and each child's body composition. Not only that, but drug doses should be adjusted for alterations in acute physiology as well. Unfortunately not enough is known about pharmacology in children in general and about drug dose modification in particular. Nonetheless, the "Holy Grail" of drug dosing is to find a "pharmacological mass" on which to base drug dose calculations, which is probably close to LBM in most children.[19] A rapidly-available, easily-achievable knowledge, or an estimation, of a child's composition would be valuable to guide the ED clinician in managing drug dosing optimally and minimising the risks of medication errors (both underdosing and overdosing).

How could we estimate body composition in the ED?

In order to be able to meaningfully evaluate body composition in a manner and time-frame compatible with emergency medical management, the system would need to be comprised of point-of-care modalities only: DXA or ADP or MRI or underwater weighing, which are commonly used methods of body composition assessment, would not be suitable. It would also need to be able to be used by someone without special skills uncommonly present in the ED (such as more advanced anthropometry including skinfold thickness measurements). The modalities that may therefore be useful might include:

- Visual assessment of body habitus (simplified from techniques performed during somatotyping) based on cognitive image analysis, a checklist of features, or comparison to printed images;
- Basic anthropometry (such as mid-arm circumference, waist-hip ratio or leg-length);
- BIA or BIS techniques (with the confounder of no measured weight to input);
- Ultrasound measurement of adipose and muscle thickness (point-of-care ultrasound skills are common in specialists practising emergency medicine worldwide, including South Africa)

There is no work that has been done to date looking at a rapid method of assessing body composition in the ED, but each of these techniques has shown promise when used in another setting, to be adaptable to use in the ED. Ultrasound in particular has shown the best-reported ability to predict BF measured by DXA when compared to BIS and ADP.

How could computerised systems contribute to patient safety?

There is a great deal of evidence that incorporating computerised systems (in whatever form they may take) into medical management protocols reduces medication errors. Little research has been done on making use of point-of-care "smart" devices in the ED which can provide decision-support as well as assist with an accurate estimation of weight and body composition and as provide complete information on drug dosing, although some related work has been done in a critical care setting. Common sense suggests that these devices can be used effectively, but there is a need for more work to confirm their effectiveness in a clinical or simulated clinical environment such as may be found in the ED. There are no freely available applications on any tablet or smartphone platform that can perform these functions (to date), and only recently have any of the available applications been subjected to scientific assessment.[20]

PROJECT 1

Basic rationale: Drug doses for any particular child should ideally be individualised according to body weight, body composition, contemporaneous physiological state and the characteristics of the drug itself. Although it is has been shown that most weight-estimation systems (including the Broselow tape) predict TBW poorly, little is known about how well they might predict a body weight form that is more appropriate to drug dose calculation. There has also been little research into multiple-step systems of weight estimation (other than the PAWPER tape) such as including, for example, basic anthropometry or BIS or ultrasound measurements of fat and muscle thickness into a simple model to accurately estimate weight and body composition. While the PAWPER tape estimates weight to within 10% in about 90% of children (compared to 40 to 60% from most other weight estimation systems) this could still be improved.

Broad research questions:

- What is the association between weight estimated by the Broselow tape, the PAWPER tape when compared with TBW, LBM, IBW and (where appropriate) ABW?
- Can the inclusion of additional modalities such as basic anthropometry, bioelectrical impedance and ultrasound into a simple PAWPER system improve the accuracy of weight estimation and body composition estimation? Could such a system differentiate between variations from the weight-for-length norm because of variations in fat mass or lean mass?
- Is it possible to create an algorithm to guide weight estimation for emergency drug dose determination that is simple and generates a pharmacologically useful base-weight for calculations?

STUDY OBJECTIVES

1. To compare the weight estimation performance by the PAWPER tape and Broselow tape against actual measured weight as well as against standard representations of body composition (LBM, IBW, ABWs) in a variety of populations of children
    a. A 'standard' population of children attending the ED
    b. A selection of high and very high weight-for-length (overweight and obese) children

c. A selection of low weight-for-length children (underweight from chronic illness)
d. A selection of children with high body water content (from chronic illness)
e. A selection of children with low body water content (from dehydration)

2. To establish whether the accuracy of weight estimation by the PAWPER tape is changed by the addition of a measure of body habitus to the standard tape-estimation technique; to investigate in which types of children (with respect to body weight and composition) the weight estimation is improved or worsened; and to investigate whether the qualifications or experience of the assessor are associated with the accuracy of the body habitus assessment.
3. To use data from the anthropometric, body composition, bioelectrical impedance and ultrasound measurements to develop a modifying factor or new system or new multipart technique that may be used to increase the accuracy of weight and/or body composition estimation based on the PAWPER tape.
4. To develop a rational algorithm to guide weight estimation techniques for use in paediatric emergency medical management.
5. To obtain photographs of the study participants for use in the Project 2.

METHODS

Design: Prospective, cross-sectional descriptive study.

Sites of study: Netcare Union Hospital and Netcare Garden City Hospital.

Study population:

- Any child from the age of 1 month up to and including 12 years of age who presents to the ED at one of the study hospitals or who is an in-patient at one of the study hospitals.
    - Group A – A convenience sample of children presenting to the emergency department
    - Group B – A group of children with a body max index of >25kg/m^2.
    - Group C – A group of children with a BMI of <18kg/m^2 from a debilitating chronic illness (e.g. HIV/AIDS)
    - Group D – A group of children who are clinically fluid overloaded from a chronic medical condition (e.g. chronic renal failure).
    - Group E – A group of children who are clinically dehydrated to at least 5% of TBW.

Inclusion criteria:

1. Children whose participation will not interfere in any way with emergent medical treatment.

Exclusion criteria:

1. Height or length >150cm (the length of the tapes);
2. Failure to obtain consent;
3. Where informed consent cannot be reasonably obtained because of a child's medical condition;
4. Where the child is too sick to be weighed or emergency medical treatment precludes participation or continued participation in the study;
5. Children with clinically significant over- or under-hydration will be excluded from Group A, but may be included in Group D or E as appropriate.

Sampling: Each of the groups will be constituted from convenience samples. Group A - any infant or child (from the age of 1 month up to and including 12 years of age) who presents to the participating EDs between 1st June 2012 and 31st December 2012 may be invited to participate (where valid consent can be obtained). Sufficient candidates will be included in the study so that there will be at least 80 individuals in each of the 3 weight divisions for the breakdown analysis. Groups B to E will be recruited via specific invitation to participate from patients attending the ED, admitted to the hospital, or attending the paediatric service at the participating hospitals. Sufficient candidates will be included in the study so that there will be at least 30 individuals in each study group. Particular care will be taken to ensure that no 'hierarchical' duress is involved in patient enrolment, and that participation in the study does not interfere with medical management in any way (see Ethics section). Sample size analysis with a power of 80% to detect a 10% difference in values at a 0.05 significance level suggested a total sample size of 78 participants for the first part of the analyses; an interim power analysis will be conducted after 50 data sets have been obtained in each weight category in order to more accurately determine a final sample size, particularly for the modelling component of the study.

Measuring instruments:

- The Broselow tape will be used in a standardised technique (there is a written description printed on the tape);
- The PAWPER tape will be used in a standardised technique (there is a written description printed on the tape) which includes weight estimation modification based on body habitus;
- Basic anthropometric measurements will be obtained using a regular SABS approved polyvinyl chloride / polyethylene tape measure;
- Actual weight will then be measured on an AE Adam Model MDW-250L digital scale (that will be calibrated according to the manufacturer's instructions) with weight recorded to the nearest 0.1kg. Children will wear only underwear and a hospital gown (of known mass) for the weighing procedure;
- BIS will be performed using a Xitron 4000B Multi-Frequency Bio-Impedance Analyzer in a standardised manner according to the manufacturer's instructions, with the child lying supine with the limbs abducted. Calibration will be performed according to the manufacturer's instructions;
- B-mode ultrasound measurements of fat thickness will be obtained using a Toshiba Nemio with a linear 5-10MHz or curvilinear 4-6MHz transducer, as appropriate, according to the methods described in Toomey et al;[21]
- DXA (DXA) will be used as a "gold standard" for body composition assessment. Whole-body DXA measurements will be obtained with paediatric software provided by the manufacturer. Calibration will be performed according to the manufacturer's instructions.

Data collection: Data will be collected by the researcher (DXA will be performed by appropriately trained radiographers). The information will be recorded on the appended data collection sheet, or stored in the memory of the electronic measurement devices and downloaded onto a computer daily. Data collection will be performed either during the triage process in a private cubicle, or after the child has been examined by the attending doctor, has received initial treatment and the weight is required to prescribe medication. The following steps will be followed:

1. The parent or guardian will be informed about the study and invited to participate – an information form, consent form and assent form, where appropriate, (see Appendix) will be given to the parent or guardian to read and complete.
2. The child will be undressed and dressed in a gown (routine in the ED); the weight of the gown will be measured; if this would not be part of the routine (e.g. for a patient with a sprained ankle) it will be specifically mentioned at the time of obtaining consent that part of the study will require the child to be dressed in a gown.
3. The data collection sheet will be initialled to confirm that consent has been obtained, and the signed informed consent document will be appended.
4. The date of data collection will be entered.
5. The child's age in years and months (e.g. 5 years 4 months) will be recorded.
6. The gender of the child (male or female) will be entered.
7. The child will be positioned supine on the examination couch for the weight estimation by the Broselow and PAWPER tapes
8. The weight of the child as predicted by the Broselow tape will be recorded.
9. The weight of the child as predicted by the PAWPER tape **without** and **with** correction for body habitus will be recorded.
10. The child will then be weighed on a digital scale.
11. The anthropometric, ultrasound and bioelectrical impedance measurements will be obtained with the child in a supine position and recorded on the data collection form, or electronic memory of the device.
12. The child (in underwear) will have a digital photograph taken against a calibrated background in a supine and erect position. They will be given a toy mask to wear to obscure their identity and only a reference number will appear on the photograph.
13. DXA measurements will be obtained at a time prior to the child's discharge – the child will be asked not to eat or urinate between the BIS and DXA measurements (a period of time which will not exceed 30 minutes).
14. The order of measurements in points 7 to 9 will be performed in an opportunistic manner and may be varied, as long as the tape weight estimations are performed before weight is measured.
15. Every 10th child enrolled will be measured twice (excluding the DXA and photographs), by different data collectors. The data will be recorded on separate collection forms. This data will be used to determine inter-observer variability.

Source of bias: This study will evaluate only children seen at a private hospital, and the results may not be generalisable to any other South African population. These children may have a different weight-for-length distribution than another population. The nursing staff will measure weights only after tape-estimated weights are recorded to avoid bias; if the weight is recorded before the measurements then this will be recorded on the data collection sheet.

Ethics: Written informed consent will be obtained from a parent or legal guardian prior to enrolment in the study (see attached consent form). Written assent will be obtained from children over the age of 7 (see

attached assent form). The parent or guardian of any child invited to participate will be given an information sheet (see attached) which includes the following points: there is no obligation to participate; participation or non-participation will in no way affect subsequent treatment; there are minimal risks, no pain or discomfort and some clear benefits to any individual patient involved in participation in the study; and efforts will be made to avoid any possible embarrassment of participants: older children will be allowed to undress and dress in a hospital gown in privacy; photographs will be taken discretely. Permission to conduct the study will be obtained from the Management of Netcare Garden City Hospital and Netcare Union Hospital. Furthermore, clearance will be obtained from the Human Research Ethics Committee for Research of the University of the Witwatersrand, and from the Netcare Research Ethics Committee.

DATA ANALYSIS

All data will be captured from the data collection forms and entered into an electronic database (Microsoft Excel, Microsoft Office 2010, Microsoft Corporation). Statistical analysis will be performed with StatSoft, Inc. (2011) STATISTICA (data analysis software system), version 10. www.statsoft.com. The data will be examined using standard parametric, non-parametric and categorical statistics to perform the following analyses:

- Evaluation of the accuracy (bias and precision) of the PAWPER and Broselow tapes in the study groups, when compared to measured weight, using paired continuous variable methods (paired t-test or Wilcoxon matched pairs test depending on data distribution) and categorical analysis (error category analysis using Fisher's exact or McNemar test).
- The actual assigned habitus score will be compared against a "perfect" habitus score (one which would have resulted in the most accurate weight estimation for that individual) and this comparison will be represented graphically as well as analysed using the percent agreement and Cohen's κ-statistic methods.
- Evaluation of the association between the tape-estimated weights and measures of body composition – specifically BMI, IBW, LBM and ABW (as above). This will allow the description
- Evaluation of the association between basic anthropometric variables and TBW and measures of body composition.
- Evaluation of the association between basic ultrasound measurements and TBW and measures of body composition.
- Evaluation of the association between bioelectrical impedance estimations of body composition with the weight and height data derived from the PAWPER tape (as opposed to measured values) and actual body composition.
- Development and evaluation of simple, resuscitation-aid compliant weight-estimation models based on the PAWPER tape, by including variables identified by regression analysis as contributing significantly to the variance between estimated and measured weight, and estimated weight and IBW and LBM. Dependent variables that will be evaluated will include: anthropometric and derived variables, ultrasound variables and raw bioelectrical impedance variables.
- Quantification of inter-observer reliability using the percent agreement and Cohen's κ-statistic methods.

PROJECT 2

Basic rationale: From the original research on the PAWPER tape, it was clear that although the difference in the design of the tape, when compared to the Broselow tape, resulted in a small improvement in the accuracy of weight estimation, it was the use of the habitus score that accounted for the majority of the dramatic improvement in accuracy. As it stands, the habitus score is assessed by a gestalt impression based on cognitive image analysis (which proved very accurate) but it is possible that a more objective method of assessing the habitus score might be more accurate, repeatable, reliable and resilient. This may take the form of a checklist or sample image outlines. Somatotyping has taught us that is it possible to quasi-quantitatively evaluate both LBM and FM with simple anthropometry or even inspection. If an observer could rapidly and reliably make an assessment of body composition (with the end-point of individualising drug dose calculation) then this would be of great benefit.

Broad research questions

- Is it possible to create a simple, easy-to-use checklist of objective morphological features that will enable an examiner to accurately determine the habitus score? Could this be more accurate than the currently used gestalt method? Could this differentiate between children who have a high BMI from increased FM, or from increased LBM? Are more than 5 categories required to represent the upper end of the weight-for-length spectrum?
- Which is the most accurate method of assessing body habitus: the gestalt method, a checklist method, or a method based on graphical representation of body habitus?
- Does assessment of the body habitus differ between healthcare providers from different disciplines and different degrees of experience?

STUDY OBJECTIVES

1. To create an objective list of criteria that will enable an examiner to accurately, consistently and reliably classify a child into a correct habitus score category; and to create a visually-based method of assessing body habitus by means of a method using drawn figures similar to those adapted by Collins[22] from the original Stunkard figures used by psychologists for self-assessment of body image (see Appendix) but with the inclusion of figures with a higher degree of mesomorphy rather than endomorphy in the high weight-for-length categories.
2. To compare the accuracy, consistency, reliability and resilience of three methods of assessing body habitus and assigning a body habitus score: the traditional gestalt method, the new checklist method and the new image-based method (see objective 1).
3. To establish whether 5 categories of body habitus are sufficient to provide optimum weight estimation.
4. To compare the abilities of different types of healthcare professionals to correctly assess body habitus using these three methods.
5. To establish the ability of healthcare professionals to guess BMI and to evaluate what methods they might use to make this “guesstimation”.

6. To develop, if appropriate, a new version of or methodology for the PAWPER tape based on the findings of this study.

METHODS

Model development

Checklist-system of body habitus assessment

The following steps will be used to derive a model to assign a body habitus score based on objective, observable criteria:

1. Potentially discriminatory criteria (between body habitus scores) that can visually and rapidly be recognised will be identified from animal condition scoring systems, adapted from photoscopic and visual somatotyping techniques and identified from the medical literature.
2. The presence, absence, or extent of these features will be ascertained from each of the children in the photographs taken in Project 1 and cross-referenced against the perfect habitus score (one which would have resulted in the most accurate weight estimation for that individual; additional habitus score categories may be added if shown to be necessary).
3. Statistical modelling, including linear and non-linear regression and principal components analysis will be used to identify the variables accounting for the greatest variance, and these will be used to create a simple matrix or checklist of criteria.
4. The derived simplified model will then be re-tested against this population to establish its accuracy.

Image-system of body habitus assessment

The following steps will be used to derive a model to assign a body habitus score based on a visual figure-based system:

1. The images adapted by Collins from the original Stunkard figures will be used as the basis for this model. These images will be adapted to add an additional series of high LBM individuals to the above-average weight-for-length children. These will be drawn up by a skilled artist.
2. These images will be used to assign a habitus score to each of the children in the photographs taken in Project 1 and cross-referenced against the perfect habitus score (one which would have resulted in the most accurate weight estimation for that individual; additional habitus score categories may be added if shown to be necessary). Information from the body composition analysis of each child will be used to quantify above average weight-for-length children as having a predominant excess of either lean or FM.
3. This data will be analysed to assess the agreement and accuracy of the system; specifically the value of having half-scores (e.g. a habitus score of 4½) and the ability of the system to discriminate between and excess fat and LBM will be evaluated.
4. Any identifiable weaknesses or inaccuracies of the system will be changed before creating the final model.

Model validation

Design: Prospective, descriptive study.

Site of study: Netcare Union Hospital and the University of the Witwatersrand Medical School (room 5A04).

Study population:

- Group A – doctors in fulltime ED practice or training;

- Group B – paediatricians or paediatric registrars;
- Group C – professional nurses;
- Group D – other categories of nurses;
- Group E – advanced life-support paramedical personnel.

Inclusion criteria: any potential member of the above group who volunteers to participate in the study.
Exclusion criteria: failure to obtain consent; previous participation in the study.

Sampling: Each of the groups will be constituted from volunteers who attend one of six continuing profession update (CPD) lectures. These lectures will address relevant issues in paediatric emergency medical care and the study will be conducted immediately after each lecture. Sufficient candidates will be included in the study so that there will be at least 30 individuals in each study group (powered to 80% at a 0.05 significance level to detect a 10% difference in scores). Particular care will be taken to ensure that no 'hierarchical' duress is involved in candidate enrolment (see Ethics section).

Measuring instruments:

- The standard gestalt-based habitus score system;
- The newly created checklist-based habitus score system;
- The newly created image-based habitus score system.

Data collection: Data will be collected following CPD lectures at six events. Attendees will be invited to participate in the study (it will be publicised during the advertising for the CPD event) which will take approximately 45 minutes. The following steps will be followed:

1. Candidates will be invited to participate and will be given an information sheet (see Appendix) and a Consent Form (see Appendix) to complete.
2. Participants will then be given the material required for the study (see Appendix), will watch a training video which will explain the organisation of the study as well as the techniques for assessing body habitus and will have an opportunity to ask questions.
3. Part 1 - The participants will then be asked to complete Section 1 of the questionnaire and will be given sufficient time so to do.
4. Part 2 – The participants will be asked to assess a habitus score from the photographs of 120 unidentifiable children and infants projected onto a screen. These photographs will have been selected from Project 1 and will be representative of the full spectrum of body compositions. The participants will have a short period of time to examine the image and record their assessment of the habitus score on an answer sheet. The first 30 cases will be assessed using the gestalt method and the allocated time will be 10 seconds per image; the middle 30 cases will be assessed using the checklist method and 30 seconds per image will be allocated; the final 30 cases will be assessed using the image method and 15 seconds per image will be allowed.
5. Part 3 – The participants will be asked to guess the BMI from the photographs of 10 unidentifiable children and infants projected onto a screen. These photographs will have been selected from Project 1 and will be representative of the full spectrum of body compositions. The participants will then be asked to write comments on what thinking processes they used to make the guesses.

Source of bias: The sequence of assessment tools for the habitus scores is designed to prevent learning from influencing the accuracy of the gestalt method. This may, however, bias the findings against this system. Similarly, fatigue may bias the findings against the latter techniques.

Ethics: Written informed consent will be obtained from the candidates prior to enrolment in the study (see attached consent form). The candidates invited to participate will be given an information sheet (see attached) which includes the following points: there is no obligation to participate; participation will be anonymous; and the participant may withdraw at any stage of the study. Permission to conduct the study will be obtained from the Management of Netcare Garden City Hospital and Netcare Union Hospital. Furthermore, clearance will be obtained from the Human Research Ethics Committee for Research of the University of the Witwatersrand, and from the Netcare Research Ethics Committee.

DATA ANALYSIS

All data will be captured from the data collection forms and entered into an electronic database (Microsoft Excel, Microsoft Office 2010, Microsoft Corporation). Statistical analysis will be performed with StatSoft, Inc. (2011) STATISTICA (data analysis software system), version 10. www.statsoft.com. The data will be examined using standard parametric, non-parametric and categorical statistics to perform the following analyses:

- Description of the demographic characteristics of the study participants (percentages for frequencies; means and 95% confidence intervals for parametric data and medians and interquartile ranges for non-parametric data);
- Evaluation of the accuracy of each of the three systems for the entire population and by subgroup (percent agreement and Cohen's κ-statistic methods, the Fisher's exact or McNemar test will be used for comparisons of categorical data);
- Comparison of the accuracy of habitus score assessment between the subgroups (the Fisher's exact test will be used for comparisons of categorical data; parametric data generated by the comparisons will be analysed using ANOVA);
- Evaluation of the accuracy of estimation of BMI (the Fisher's exact test will be used for the analyses);
- Qualitative assessment of the comments made with regard to the method of BMI "guesstimation" (standard qualitative evaluation techniques will be used).

PROJECT 3

Basic rationale: Virtually every organisation across the world that has examined the issue of medication errors in children has made the recommendation that information technology be incorporated into the clinical management strategy. This may, at the lower end, be simple computer-based prescription systems or, at the upper end, full decision-support software systems. The use of portable point-of-care devices such as smartphones, tablets or other devices for paediatric drug dose guidance has become popular, but is largely untested, and no applications exist that incorporate length-based weight estimation, or drug dose

modification based on body composition. Such an application would be a useful tool in the hands of a team providing emergency medical care to children.

Broad research questions:

- Is it possible to develop an application for the Apple iPad and Apple iPhone that can function as a comprehensive, reliable, resilient resuscitation aid that incorporates a system of length-based, habitus-modified weight estimation and provides full drug dosing (and equipment-sizing) information that is modified according to body composition (when appropriate)?

STUDY OBJECTIVES

1. To use the information derived from Projects 1 and 2 to generate the theoretical background for an Apple application that will:
    a. Be able to accept input of actual measured weight; or weight estimated by the PAWPER tape or Broselow tape or any of the other currently available tapes; or measured length;
    b. Be able to rapidly guide the user through the process of habitus score assignment (based on the findings of Project 2);
    c. Be able to provide equipment size selection guidance (which is best predicted by body length and is relatively independent of actual weight);
    d. Be able to provide comprehensive information on drug dose and drug preparation for any parenteral drug that might commonly be used in the ED, modified according to individual drug, body composition and acute physiological state;
    e. Be compliant with the principles of a resuscitation aid and the principles of the weight-estimation algorithm defined in Project 1.
2. To test the developed application against the PAWPER tape + Emergency Drug Dosing in Children resuscitation aid book, the Broselow tape and a wild card system in a series of simulated clinical scenarios.

METHODS

Application development: The application will be developed by an expert under the direction of the researcher, based on the findings of Projects 1 and 2, and incorporating drug dosing information already developed by the researcher.

Application testing

Design: Prospective, cross-over study.

Sites of study: University of the Witwatersrand Medical School.

Study population: Emergency Medicine consultants and Emergency Medicine registars, in any year of study, affiliated with the Division of Emergency Medicine, Faculty of Health Sciences, University of the Witwatersrand.

Inclusion criteria: any member of the study population pool.

Exclusion criteria: failure to obtain consent.

Sampling: The study sample will be constituted from volunteers from the consultant, registrar and academic staff from Division of Emergency Medicine, Faculty of Health Sciences, University of the Witwatersrand. At

least 30 participants will be enrolled (powered to 80% at a 0.05 significance level to detect a 10% difference in scores).

Measuring instruments:

1. The PAWPER tape plus the Emergency Drug Dosing in Children book will constitute the first resuscitation aid (PAWPER + EDDC);
2. The application developed in Project 3 will constitute the second resuscitation aid (PAWPER APP)
3. The Broselow tape will constitute the third resuscitation aid (it was the first and still most commonly used non-computerised resuscitation aid).
4. The participants' own choice of other weight estimation method, without any resuscitation aid for drug dosing, will act as the control group.

Data collection: Two study days will be selected, to each of which 15 volunteers will be invited. Eight stations will be set up, each with a child volunteer as a patient, with a variety of simulated common emergency situations. Each participant will conduct an abridged simulated resuscitation at each station as directed by an instructor (data-collector) which will involve determining the appropriate administration of two drugs and appropriate selection of one piece of equipment; two of the stations will have the PAWPER + EDDC, two will have the PAWPER APP; two will have the Broselow tape to assist the participant; and two will have no resuscitation aid. Each station will take no longer than five minutes to complete. The study will be conducted in the following way:

1. At the beginning of the study day, the volunteers will be given an information sheet (see Appendix) and a Consent Form (see Appendix) to complete.
2. Participants will then watch a training video which will explain the organisation of the study as well as the techniques for using the tapes, the PAWPER APP and for assessing body habitus. They will have an opportunity to ask questions and to completely familiarise themselves with the resuscitation aids in their own time.
3. Once all the participants are ready, the study will commence. Each participant will complete each of the eight stations using each system for two stations – the sequence of device usage will be according to a predetermined table which will ensure that each child is assessed an equal number of times with each device and in as random a pattern as possible.
4. At each station the participant will have four tasks to perform: weight estimation (by the PAWPER tape, the PAWPER APP, the Broselow tape and the participant's chosen method); two drug dose calculations to perform, and one piece of equipment to correctly size. These answers will be written on a proforma data collection sheet. The instructor will explain each task, give the participant the instruction to begin, and then record how long it takes them to complete the task, incomplete tasks, any errors that were made as well as any comments on difficulties the participants experienced.
5. At the end of the three tasks (or after a maximum of 5 minutes), the participant will move to the next station.
6. At the end of the cycle of six stations, participants will be invited to make any written comments about the three resuscitation aid systems;
7. The children 'models' will be volunteers and will only be used for weight-estimation purposes and no actual or simulated procedures. They will have their TBW measured, their length measured by an experienced instructor which will be regarded as the "reference" length, and their body composition

will be assessed using BIS. They will be selected so that a spectrum of body habitus types is represented.

Source of Bias: The learning effect might cause the participants' performance in the latter stations to exceed that in the former.

Ethics: Written informed consent will be obtained from the candidates prior to enrolment in the study (see attached consent form). The candidates invited to participate will be given an information sheet (see attached) which includes the following points: there is no obligation to participate; participation will be anonymous; and the participant may withdraw at any stage of the study. Permission to conduct the study will be obtained from the Human Research Ethics Committee for Research of the University of the Witwatersrand.

DATA ANALYSIS

All data will be captured from the data collection forms and entered into an electronic database (Microsoft Excel, Microsoft Office 2010, Microsoft Corporation). Statistical analysis will be performed with StatSoft, Inc. (2011) STATISTICA (data analysis software system), version 10. www.statsoft.com. The data will be examined using standard parametric, non-parametric and categorical statistics to perform the following analyses:

- Description of the demographic characteristics of the study participants (percentages for frequencies; means and 95% confidence intervals for parametric data and medians and interquartile ranges for non-parametric data);
- The performance data for each of the resuscitation aids will be determined as follows (paired categorical data will be analysed using the McNemar test, unpaired with the Fisher's exact test; continuous paired data will be compared with the paired t-test or Wilcoxon matched pairs test, unpaired with the t-test or Mann-Whitney test; analysis of multiple samples with be performed with ANOVA):
    - The number and magnitude of any errors;
    - The number of incomplete tasks;
    - An average or median time (distribution dependent) for the completion of the medication dose tasks;
    - An average or median time (distribution dependent) for the completion of the equipment tasks;
    - Any comments by the instructors or participants will be evaluated using qualitative analytical techniques
- The drug doses that were determined for each child will be compared to an "ideal" dose that takes body composition and acute physiological state into account – the nature and magnitude of any differences will be analysed as above.

TIMING

The Gantt chart shows the approximate timing for the various projects and procedures.

	2012												2013											
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
Ethics																								
Project 1																								
Project 2 model development																								
Project 2																								
Project 3 model development																								
Project 3																								
Data analysis																								
Write-up																								
Submission																								

FUNDING

Any incidental expenses will be borne by the principal investigator; funding will be sought from the Medical Research Council and the National Research Council.

POTENTIAL PROBLEMS

1. If DXA proves to be too expensive (in the absence of adequate funding), too time-consuming or not available, then BIS will be substituted as the method of choice for assessing body composition.

REFERENCES

[1] Kozer E, Berkovitch M, Koren G. Medication errors in children. Pediatr Clin N Am. 2006;53:1155-1168.

[2] System. CotFoECitUSH. Emergency Care for Children: Growing Pains. IOM, editor. Washington, DC: National Academy Press; 2006.

[3] Nash DB. Meet the enemy: he is us. American Journal of Medical Quality. 2011;26(1 Suppl):4S.

[4] Fein IA. Quality, patient safety, and culture: 'We have met the enemy and he is us'--Pogo (Walt Kelly, 1971). Critical Care Medicine. 2011;39(5):1196-7.

[5] Barata IA, Benjamin LS, Mace SE, Herman MI, Goldman RD. Paediatric patient safety in the prehospital/emergency department setting. Ped Emerg Care. 2007;23(6):412-418.

[6] Luten R. Error and time delay in pediatric trauma resuscitation: addressing the problem with color-coded resuscitation aids. Surg Clin North Am. 2002;82:303-314.

[7] Luten R, Wears RL, Broselow J, Croskerry P, Joseph MM, Frush K. Managing the unique size-related issues of paediatric resuscitation: reducing cognitive load with resuscitation aids. Acad Emerg Med. 2002;9:840-847.

[8] Conroy S, Sweis D, Planner C, Yeung V, Collier J, Haines L*, et al.* Interventions to reduce dosing errors in children. Drug Safety. 2007;30(12):1111-1125.

[9] Wong ICK, Ghaleb MA, Franklin BD, Barber N. Incidence and nature of dosing errors in paediatric medications. Drug Safety. 2004;27(9):661-670.

[10] Mack A. Policy statement--principles of pediatric patient safety: reducing harm due to medical care. Pediatrics. 2011;127(6):1199-210.

[11] Cowley E, Williams R, Cousins D. Medication errors in children: a descriptive summary of medication error reports submitted to the United States Pharmacopeia. Current Therapeutic Research, Clinical and Experimental. 2001;62(9):627-640.

[12] Lesar TS, Mitchell A, Sommo P. Medication safety in critically ill children. Clin Ped Emerg Med. 2006;7:215-225.

[13] Croskerry P, Shapiro M, Campbell S, *et al.* Profiles in patient safety: medication errors in the emergency department. Acad Emerg Med. 2004;11:289-299.

[14] Fineberg SL, Arendts G. Comparison of two methods of paediatric resuscitation and critical care management. Ann Emerg Med. 2008;52:35-40.

[15] Garland JS, Kishaba RG, Nelson DB, Losek JD, Sobocinski KA. A rapid and accurate method of estimating body weight. Am J Emerg Med. 1986;4:390-393.

[16] Wells M. Weight prediction in children in the emergency department. Johannesburg: University of the Witwatersrand; 2009. Available from:

[17] Losek JD, Garland J, Nelson DB. Body habitus + height = accurate weight estimate. Ann Emerg Med. 1989;18(6):709.

[18] Yamamoto LG, Inaba AS, Young LL, Anderson KM. Improving length-based weight estimates by adding a body habitus (obesity) icon. Am J Emerg Med. 2009;27(7):810-5.

[19] Dieckmann RA. The dilemma of paediatric drug dosing and equipment sizing in the era of patient safety. Emergency Medicine Australasia. 2007;19:490-493.

[20] Flannigan C, McAloon J. Students prescribing emergency drug infusions utilising smartphones outperform consultants using BNFCs. Resuscitation. 2011;82(11):1424-7.

[21] Toomey C, McCreesh K, Leahy S, Jakeman P. Technical considerations for accurate measurement of subcutaneous adipose tissue thickness using B-mode ultrasound. Ultrasound. 2011;19:91-96.

[22] Collins M. Body figure perceptions and preferences among preadolescent children. International Journal of Eating Disorders. 1991;10(2):199-208.

[23] Wells M, Goldstein L, Botha M. Emergency drug dosing in children: a resuscitation aid for paediatric emergencies. Edinburgh: Elsevier; 2012.

APPENDIXES

LIST OF ABBREVIATIONS

ABW	Adjusted Body Weight
ADP	Air Displacement Plethysmography
BIA	Bioelectrical Impedance Analysis
BIS	Bioelectrical Impedance Spectroscopy
BMI	Body Mass Index
DXA	Dual Xray Absorptiometry
ED	Emergency Department
EDDC	Emergency Drug Dosing in Children book
FM	Fat Mass
HS	Habitus Score
IBW	Ideal Body Weight
LBM	Lean Body Mass
MRI	Magnetic Resonance Imaging
TBW	Total Body Weight

The JCAHO sentinel alert of April 2008

On April 11 2008 the Joint Commission on the Accreditation of Healthcare Organisations (JCAHO) in the USA issued an alert entitled "Preventing paediatric medication errors". This report made the following major recommendations that can be applied to paediatric prescribing in the ED:

- A paediatric formulary should be available in every institution.
- Paediatric specific reference material (including website access) should be available for doctors involved in the emergency care of children, including growth charts, normal vital signs for different age-groups of children, emergency dosage calculations and drug reference materials.
- Weight (in kilograms) must be measured in all paediatric patients at the time of presentation, or no later than 4 hours after admission in emergency cases.
- Drug prescriptions should include the dose calculation to facilitate double-checking of the prescription. For example in a 13kg child: *administer phenytoin 195mg intravenously over 20 minutes (15mg/kg/loading dose: dilute 250mg/5ml phenytoin + 45ml 5%DW and infuse 39ml)*.
- Discharge medications should be prescribed similarly (*administer amoxicillin SF 7.5ml orally three times a day (250mg/5ml at 30mg/kg/dose))* and oral syringes should be used for administration.
- Ongoing training should be provided to doctors and nurses involved in the care of children with respect to medication prescription and delivery.

The Paediatric Advanced Weight Prediction in the Emergency Room (PAWPER) tape.

This length-based weight prediction tape was devised to improve the accuracy of weight prediction in children. Studies have shown the Broselow tape to be inaccurate in some populations and specifically in children in the upper centiles for weight (and obese children). While superficially the PAWPER TAPE appears similar to the Broselow tape, the reasoning behind the construction is fairly different. In many ways it is easiest to describe its features by differentiating them from the Broselow tape (see point 2).

1. What is the PAWPER tape?

   The PAWPER tape is a full scale approximation of the 50th centile of the weight-for-length chart produced by the World Health Organisation (WHO). Each weight division corresponds to the 50th centile of length for that weight (this centile corresponds to the centre of each weight division). Since not every child has a weight on the 50th centile for their length, the 5th and 95th centiles are shown in parenthesis for each weight division. It has been shown previously that weight can be predicted more accurately when a measure of frame-size is included in the prediction methodology. The PAWPER tape is designed so that the weight of children with greater than average weight-for-length is not underestimated. This is based on common-sense – it may be necessary to create a more objective method of frame assessment, but this study should provide the answer to that question.
2. How does the PAWPER tape differ from the Broselow tape?
   a. It is a single use, disposable device, projected to be available at a cost of less than 1 US dollar.
   b. The PAWPER tape contains no information on drug doses or equipment sizes.
   c. The PAWPER tape is based on the WHO growth charts, not the American CDC 2000 growth charts.
   d. The PAWPER tape reflects the 50th centile in the centre of each weight division; the Broselow tape reflects the 50th centile at the distal edge of the weight division. In other words, the PAWPER tape would predict an actual weight of 13.8kg to be 14kg, while the Broselow tape would predict an actual weight of 13.8kg to be 13kg.
   e. The PAWPER tape includes the 5th and 95th centiles, with a method to adjust the weight according to the frame size or degree of adiposity of the patient.
3. How is the PAWPER tape similar to the Broselow tape?
   a. In order to maintain standardisation the same weight and colour divisions found in the Broselow tape have been adhered to.
4. How is the PAWPER tape used?
   a. See below – extracted from Wells et al[23] with permission.

How to use the Paediatric Advanced Weight Prediction in the Emergency Room (PAWPER) tape

1. Place child supine on the stretcher.
2. Unroll tape alongside the child (Fig. 1.5).
3. Measure from the child's head to their heel – placing the side that has the five arrows and says 'MEASURE FROM THIS LINE' at the child's head.
4. Take note of the weight in the block obtained.
5. The main bold weight (Fig. 1.6) can be used if the child is of average build.

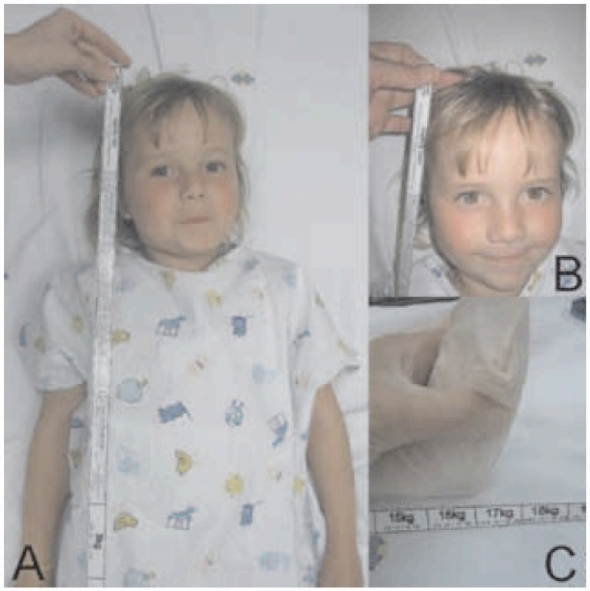


Fig 1.5 How to use the PAWPER tape. A – Unroll the tape across the front of the supine child from head to heel; the PAWPER tape is flexible and easy to align. B – Use the fingers of the hand at the patient's head to feel for the vertex to align the end of the tape with the top of the head. C – Flex the foot to 90° and note where it crosses the tape; it is essential to remove shoes and ideal to remove socks or stockings. This might be easier with a second person in a large child.



Fig 1.6 A segment of the PAWPER tape showing the estimated weights for each of the habitus score categories – from left to right: skinny (HS1), slightly thin (HS2), average (HS3), slightly chubby (HS4), fat (HS5).

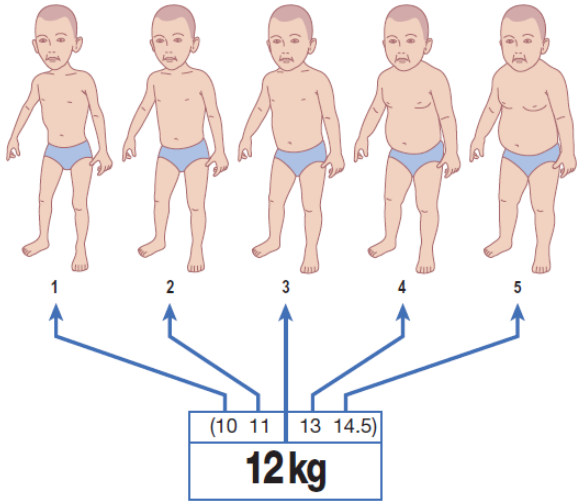
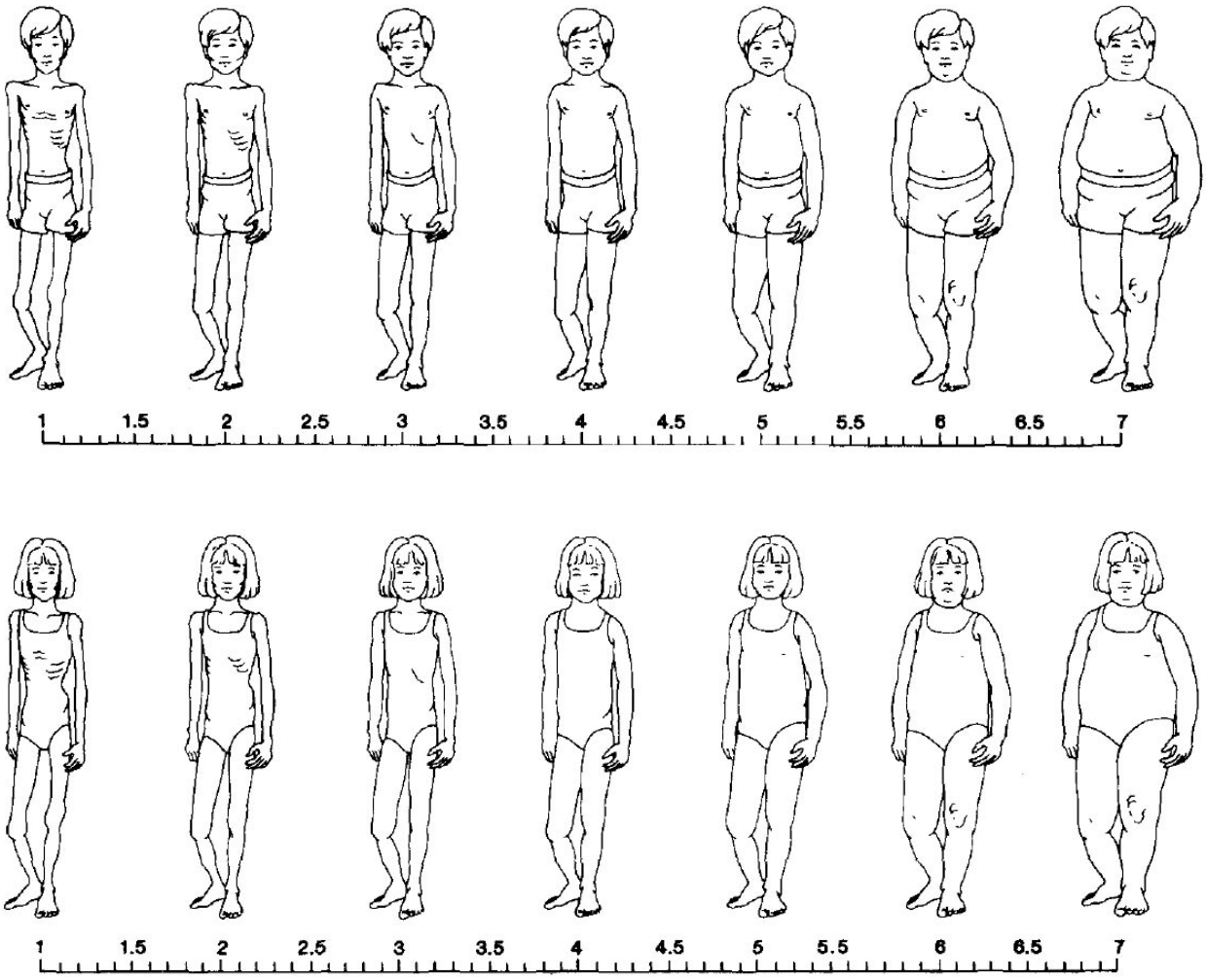


Fig 1.7 Differing weight-for-length in children. The use of five habitus scores with corresponding weights in each segment of the PAWPER tape allow for a more accurate weight estimation than that generated by other tape systems. A habitus score of 1 corresponds to a weight at the 5th centile of weight-for-length, a score of 2 with the 25th centile, a score of 3 (average) with the 50th centile, a score of 4 with the 75th centile and a score of 5 with the 95th centile. It is possible using a rapid 'eyeball' technique to accurately estimate deviation of body habitus from the average and to adjust the weight estimation accordingly.

6. The other four weights are used if the child is not of average build in order to adjust for the child's body habitus. These weights correspond to the habitus score (HS) which is assigned during the tape measurement procedure (Fig. 1.7).
7. Check for the drug required and the dosage thereof in this book
8. If the drug required is not available, use the weight measurement obtained in order to calculate the dosage.



The Collins figure rating scales for boys and girls – as used for the self-assessment of body image.

Appendix 2
Ethics clearance certificates

The clearance certificates from the Human Research Ethics Committee of the University of the Witwatersrand are attached. The first certificate provides clearance for the three prospective studies included in the thesis. The second certificate provides clearance for the secondary data analysis of three previous studies (each of which had individual ethics clearances).

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG
Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Dr Mike Wells

CLEARANCE CERTIFICATE	**M120486**
PROJECT	Estimating Children's Weight in the Emergency Department using a Length-Based, Habitus-Modified System to Increase the Accuracy of Total Body Weight Estimation: The Utility of the PAWPER Tape in Paediatric Emergency....
INVESTIGATORS	Dr Mike Wells.
DEPARTMENT	Division of Emergency Medicine
DATE CONSIDERED	04/05/2012
DECISION OF THE COMMITTEE*	Approved unconditionally

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

DATE 04/05/2012 **CHAIRPERSON** .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable
cc: Supervisor : Dr Alison Bentley

DECLARATION OF INVESTIGATOR(S)
To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...

R14/49 Dr Mike Wells

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M1511107

NAME: (Principal Investigator) Dr Mike Wells

DEPARTMENT: Emergency Medicine

PROJECT TITLE: Advances in Paediatric Emergency Weight Estimation - Secondary Analysis of Prospectively Collected Data (from previously approved protocols M080413, M120486 and M140510)

DATE CONSIDERED: Adhoc

DECISION: Approved unconditionally

CONDITIONS:

SUPERVISOR:

APPROVED BY: Professor A Dhai, Co-Chairperson, HREC (Medical)

DATE OF APPROVAL: 15/12/2015

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

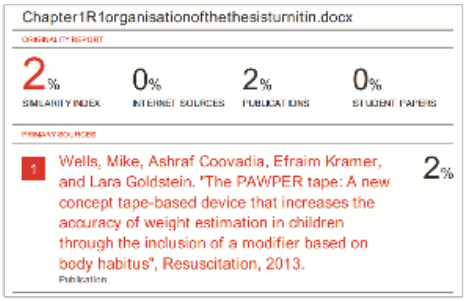
To be completed in duplicate and **ONE COPY** returned to the Secretary in Room 10004, 10th floor, Senate House, University.
I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report**.

Principal Investigator Signature Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Appendix 3
Turnitin report

The Turnitin reports generated for this thesis are attached.



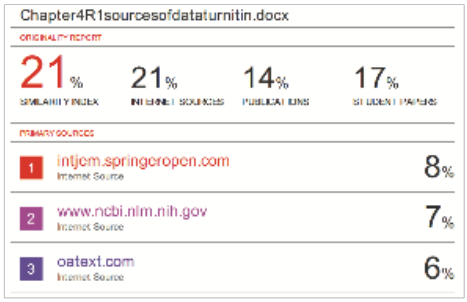
Chapter1R1organisationofthethesisturnitin.docx

ORIGINALITY REPORT

2%	0%	2%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Wells, Mike, Ashraf Coovadia, Efraim Kramer, and Lara Goldstein. "The PAWPER tape: A new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus", Resuscitation, 2013. Publication	2%



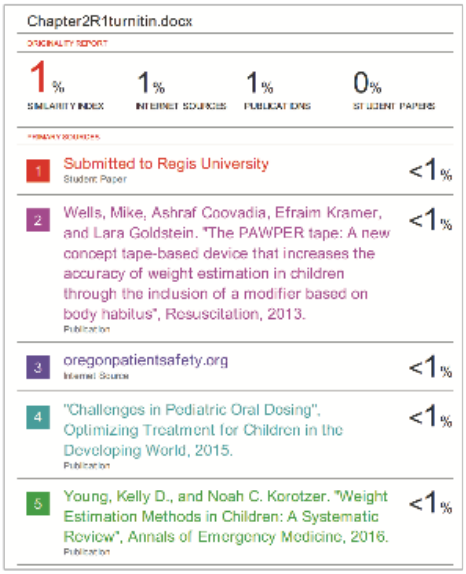
Chapter4R1sourcesofdataturnitin.docx

ORIGINALITY REPORT

21%	21%	14%	17%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	intjem.springeropen.com Internet Source	8%
2	www.ncbi.nlm.nih.gov Internet Source	7%
3	oatext.com Internet Source	6%



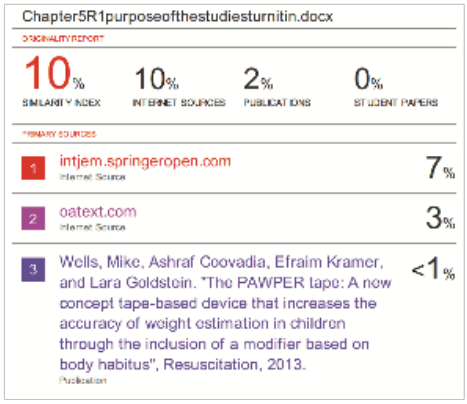
Chapter2R1turnitin.docx

ORIGINALITY REPORT

1%	1%	1%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Regis University Student Paper	<1%
2	Wells, Mike, Ashraf Coovadia, Efraim Kramer, and Lara Goldstein. "The PAWPER tape: A new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus", Resuscitation, 2013. Publication	<1%
3	oregonpatientsafety.org Internet Source	<1%
4	"Challenges in Pediatric Oral Dosing", Optimizing Treatment for Children in the Developing World, 2015. Publication	<1%
5	Young, Kelly D., and Noah C. Korotzer. "Weight Estimation Methods in Children: A Systematic Review", Annals of Emergency Medicine, 2016. Publication	<1%



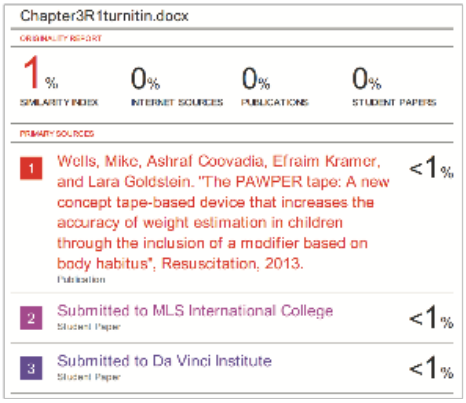
Chapter5R1purposeofthestudiesturnitin.docx

ORIGINALITY REPORT

10%	10%	2%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	intjem.springeropen.com Internet Source	7%
2	oatext.com Internet Source	3%
3	Wells, Mike, Ashraf Coovadia, Efraim Kramer, and Lara Goldstein. "The PAWPER tape: A new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus", Resuscitation, 2013. Publication	<1%



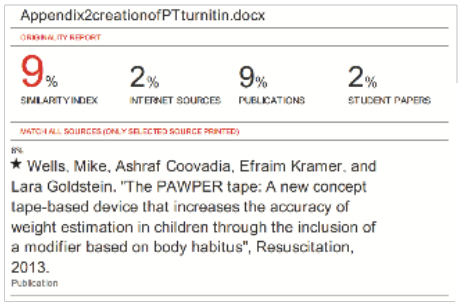
Chapter3R1turnitin.docx

ORIGINALITY REPORT

1%	0%	0%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Wells, Mike, Ashraf Coovadia, Efraim Kramer, and Lara Goldstein. "The PAWPER tape: A new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus", Resuscitation, 2013. Publication	<1%
2	Submitted to MLS International College Student Paper	<1%
3	Submitted to Da Vinci Institute Student Paper	<1%



Appendix2creationofPTturnitin.docx

ORIGINALITY REPORT

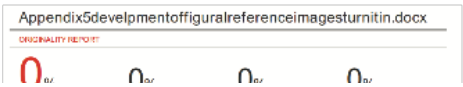
9%	2%	9%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

8%

★ Wells, Mike, Ashraf Coovadia, Efraim Kramer, and Lara Goldstein. "The PAWPER tape: A new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus", Resuscitation, 2013.

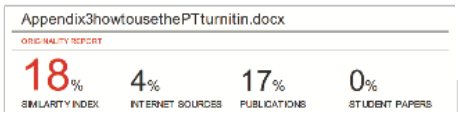
Publication



Appendix5develpmentoffiguralreferenceimagesturnitin.docx

ORIGINALITY REPORT

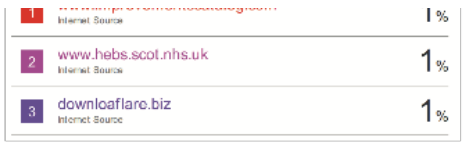
0%	0%	0%	0%



Appendix3howtousethePTturnitin.docx

ORIGINALITY REPORT

18%	4%	17%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS



1	Internet Source	1%
2	www.hebs.scot.nhs.uk Internet Source	1%
3	downloaflare.biz Internet Source	1%



Appendix7forEDD4kturnitin.docx

ORIGINALITY REPORT

4%	0%	4%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Imms, Christine. "Review of the Children's Assessment of Participation and Enjoyment and the Preferences for Activity of Children", Physical & Occupational Therapy In Pediatrics, 2008. Publication	4%