DEMAND MANAGEMENT IN HEALTH CARE:

THE CASE FOR FAILURE DEMAND

Dieter Hartmann

A thesis submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

Johannesburg 2021

Declaration

I declare that this thesis is my unaided work. It is being submitted for the Degree of Doctor of Philosophy to the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination to any other University.

Signed at Johannesburg, South Africa

-------------------------------------- 01/01/2021

Dieter Hartmann

Abstract

Failure demand has been shown to have a material impact in many service industries – leading to increased waiting times and reduced system capability. The nature and impact of failure demand in health systems has however not been studied in great depth.

This study proposes managing demand, and more finely, failure demand as an alternative focus for closing the gap between capacity and demand. This is contrasted against the traditional focuses on system capacity, which is raised through investment or efficiency improvements.

To manage demand, the context must be understood, so a definition of the demand population for the health system is proposed, out of which a proposal is made for a mental model that describes the demand-modalities that exist in health systems. This model contains four key demand classes, namely, value- and failure - demand (using Seddon's terminology) and expanded by adding escalation demand and false demand. Failure demand is selected for development and an algorithm is proposed that defines failure demand in a complex hierarchical organisation such as health care. A table is presented of common events that drive failure demand in health care. Leading out of this model, a health care setting is selected, in this case, a national pharmaceutical supply-chain in a developing country.

The analysis was conducted by data mining order- and dispatch-documents and virtually recreating the operating history. For this, custom code was developed in Visual Basic for Applications, using a Sequential Pattern Mining approach. The Wholesale- and Distribution-networks were analysed and failure demand levels of 56 % and 29 % respectively were found in these networks. Significant service delivery improvements are foreseen if the root causes of failure demand are addressed, which in this case are mainly procurement-policy related.

The study shows that failure demand in health systems represents an opportunity to narrow the capacity-demand gap by managing demand through targeted interventions.

Acknowledgements

Acknowledging the people who made this PhD possible is impossible. Regretfully, I will forget or diminish the profound role that was played by so very many people.

Thank you to Bruno, a wise, thorough and supportive supervisor. Who would catch every error, and supply deep and insightful feedback, and for your friendship in tough and easy times.

To John – a Lean guru – the most well-read human in Lean and perhaps in all topics of human endeavour. John has been sharper than a knife, always supportive, and my mentor in this exciting field of study.

To Teresa, you shaped this study, and you should have had your name listed as a supervisor. But more so, for repeatedly rescuing this doctorate, and also - me.

To Sharon, I fear few people more than you, but you have been a believer and a rock-solid supporter from the very first day I met you almost a decade ago. Without you, I can assure you that this PhD would never have been handed in – possibly even started, let alone restarted several times.

To little Max. I am sorry that I had to go on so many trips to get this work done. I am sorry that you grew up so much with me sitting in Kampala and Ibadan and other places whilst you learned to walk and talk and grow up.

To Carol, this interest started with you; I read the realisation in the rings under your eyes. The realisation that health care could be more efficient and less harmful to all who participate in it. And although this work may not fix much, I believe it speaks to the idea that change and improvement are possible. Thank you for your support the space and allowing me to leave home for weeks at a time, without complaining, to get this doctorate done.

To CARTA – fellows and teachers and administrators and an all-around team of exceptionalism. Your support has groomed me as a scholar more than any other. Attending the JASes in beautiful cities across the African continent has opened my eyes and given me enthusiasm for the places and ways in which health can come to more people, better and easier.

To Alex, for beautiful insights and thoughts. For introducing me to the idea that the unsatisfied burden of poor healthcare shifts to society.

To Opa – a true citizen scholar, a man who recited Gilgamesh – in Akkadian and taught evolution when it was illegal, you debated with priests and physicists on causality. You knew the geological structure of South Africa like no other. You found fossils, climbed mountains and lived the raconteur. You inspired me to know stuff and to be curious about everything.

To Johann Sebastian Bach – thank you for being in my ears always. You really did get so terribly much right. SDG.

Und Mama und Papa, ich glaube, dass ich das bin was ich bin wegen euch. Ihr habt mir immer Spielraum und Freiheit gegeben, darin konnte ich frei denken, schöpfen und viele Dinge vernichten. Aber aus dem Ganzen lernte ich wie Sachen funktionieren, wie das Leben funktioniert, und was es heißt kritisch und frei zu denken. Ich hoffe, dass ein kleiner Teil von diesem in dieser These wiederzufinden ist.

Table of Contents

List of Tables

List of Figures

Nomenclature and Terminology

Capacity: The ability to deliver a product, commodity or service. May be interpreted as the design capacity, which implies full efficiency.

CDC: US Centers for Disease Control and Prevention

Demand: The load that is placed on a system, due to the needs of the customers.

Failure Demand: the load on a system due to returning instances of previous system failure.

Escalation Demand: The demand that has escalated due to delay in treatment or action.

False demand: The load placed on a health system by the healthy population.

Morbidity Demand: The demand created by the, undeniably ill, like value demand, however morbidity can also contain failure and escalation demand.

Non-health motives for demand: the demand created when patients place a load on a health system for administrative reasons, such as sick leave certificates etc.

Patient preference demand: the demand created because patients prefer to consume a service from one service provider rather than from another.

Perceived need Demand: the demand created by the belief of a patient that they need to consume a health care service.

Supply Induced Demand: the demand generated in a system by the health care provider, particularly by encouraging an uninformed patient to consume an unnecessary service, test or intervention.

Value Demand: The correct, first time and appropriate demand.

Effectiveness: A measure of the extent to which a service achieves its intended outcomes.

Efficiency: a measure of the extent to which resources input into a transformation are optimally converted to the desired output.

ERP: Enterprise Resource Planning (System)

GP: General Practitioner

Health: a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.

HMO: Health Management Organization - health insurance that allows access to networked primary health services on demand for no, or low, patient payments, variable based on patient risk profile.

JIT: Just In Time

OECD: Organization for Economic Co-Operation and Development

Supply: In economics, the gross availability of a product, commodity or service.

Supply-Demand Gap: with demand being generally greater than supply, leading to an inability to serve all patients in a satisfactory manner.

Variation: the absence of stability, in this case particularly, the absence of repeatable anticipatable behaviours.

Variability: represents the state in which variation takes place.

VBA: Visual Basic for Applications

WHO: World Health Organization

WIP: Work in Progress

WMS: Warehouse Management System

Chapter One: **Introduction**

This chapter has three purposes, firstly, to introduce the reader to the context and importance of this study. Secondly, some previous work that had been done that highlighted the gaps in understanding and finally to guide the reader through reading the rest of this thesis.

The concluding parts of this chapter explain how the study was structured. This includes introducing the rationale, aim and the research questions. After this, the structure of this thesis as a doctorate by publication is discussed and concludes with some suggestions on how to read this thesis.

1 - 1. Introduction

Arguing for revolutionary change in the health sector, Merry highlights the paradox that health care is the largest industry in our society, yet it is run according to the craft principles of its origins in the eighteenth century [1 Pg. 4]. Total health expenditure accounts for 9 % of all expenditure globally [2] which accounts for about 5.6 Trillion US Dollars [3] – or put another way the combined size of the economies [4] of the UK, France and Italy. The health care percentage of 9 % compares to approximately 5 % for global education spending [3, 5, 6] and less than two percent for military spending [3].

A major component of the burden of health care is the cost of pharmaceuticals [7, 8]. A recent report reveals how a Utah medical insurer pays for the flights of selected patients from Utah to San Diego, transports them across the Mexican border to Tijuana, where they fill their prescriptions, and return home. In addition to all costs for the travel, the insurer also pays the patients a stipend of US$ 500 to make use of this facility. The insurer believes that this cross-border initiative allows them to get the most expensive scripts at half the cost, and has already saved them US$ 225 000 in less than a year [9].

This leads to two questions, firstly whether health care needs to be so costly, and secondly whether there are modern ways to run the industry that would improve outcomes.

To improve and expand service delivery, the traditional approach is to make the health care system bigger and more expensive [10-12], whilst industrial engineering and Lean approaches in health care focus on improvement initiatives to make systems more efficient [13]. The latter approach results in systems that do more with less, finding capacity through efficiency, and providing a higher quality service by eliminating waste and ineffective practice [14, 15].

A central notion in this thesis is the concept of demand and understanding it. This thesis proposes that demand in health care consists of several different building blocks which together account for the load on health systems. If the modes of demand that are unnecessary can be purposefully managed, the overall workload of medical personnel may be reduced.

This thesis speaks about the gap between the capacity of a health system and contrasts that against the demand on this system. It presents the demand-capacity gap and proposes remedies for closing this gap by managing demand rather than the more common focus on managing capacity. Particular attention is paid to failure demand and understanding root causes for it. This thesis uses the Lean point of view that efficiency and elimination of waste are virtuous, and that improving these for a system achieves overall positive results.

1 - 1.1. Pilot studies that shaped this thesis.

In trying to understand failure demand, this author conducted three pilot studies, prior to embarking on the work for this thesis. United by a common thread of failure demand, these pilot studies revealed a number of weaknesses and gaps in the current theoretical environment. This thesis should in part set a foundation for future, similar studies. Some of the methodological characteristics of these approaches are also covered in Section 3 - 3 – “Abandoned methods”.

The first pilot study was conducted in the operating theatre of a major public hospital. It made use of observations and records analysis [16, 17]. One of the key realisations of this pilot study, was that the definitions and frameworks used to identify failure demand were insufficient for the medical environment. As a result, the outcomes of this pilot study, were questionable, based on the under reporting of failure demand. Major concerns emerging from this included that complex hierarchical movements were not considered in the definition, nor was there specific clarity about which behaviours reflected system failure, and which did not.

The second pilot study was conducted in a radiology department over a period of eight weeks, during which time records were analysed. The effect of poor record keeping, and data integrity showed the importance for finding a suitable data set. This study further found the importance of the patient as a part of the system – leading to the finding that system failure is equally a consequence of patients' behaviours as it is by the care-providers. This study highlighted the importance of infrastructure integrity and human resource reliability. In the absence of these – it became clear that the conclusions of what failure demand is, is not clear.

The third pilot study investigated a "big data" set from a public hospital. In this study, patient movements were data-mined through general admissions. The findings of this study, were that inadequate adherence to complete data capture, led to the inability to track patient movements through the system over time. The consequence of this, is that all findings from the study were questionable. Therefore, it was clear that any future attempt to trace movements would require rich comprehensive and correct data sets.

From the pilot studies, three key requirements emerged. Firstly, that the nature of failure demand in health systems be rigorously refined and an framework be designed that can be used to assess a practice to classify it as failure demand. Secondly, that an understanding of the demand populations is created and thirdly, that analysis requires comprehensive movement data that do not conceal failure demand in absent data points.

Even so – these pilot studies provided rich contextual knowledge about health systems, issues experienced and formed a part of the eventual mental models presented in this thesis that could bridge the gap in understanding failure demand broadly.

1 - 2. Roadmap for Reading This Thesis.

1 - 2.1. Structure of the thesis

This thesis is structured as a doctorate-by-publications - an emergent style-preference in engineering doctorates, particularly as it serves the dual mandate of knowledge- creation and -dissemination most efficiently [18]. The thesis conforms to the guidelines of the University of the Witwatersrand, Johannesburg. This thesis contains three submitted papers. The University does not require that papers be published – however at least one paper must have been submitted for publication, and all three have.

A high-level discussion of the methods used in this thesis is presented in Chapter 3. This chapter discusses the overall scientific structure and presents a thread for the argumentation followed. More detailed method, as required, is contained in the further chapters as required.

The papers are contained in Chapter 4, Chapter 5, and Chapter 6. The papers have all been written for targeted journals. The final chapter is written as an integrating narrative, to unite the argumentation from the papers into a single story.

1 - 2.2. Rationale

The rationale for this research is a belief that health care delivery can be improved if the gap between capacity and demand is narrowed. This study focuses on demand[1], and more narrowly failure demand. The study will address:

- The lack of a clear mechanism-driven structure defining the modalities composing demand,
- The absence of a clear framework that explores the complexities of the phenomenon of failure demand in health care systems,
- The low amount of technical benchmark values for the effect of failure demand on health provision.

1 - 2.3. Aim

To investigate failure demand and its role in shaping demand in health systems.

1 - 2.4. Research question:

What impact does failure demand have on the gap between capacity and demand in health systems?

1 - 2.5. Objectives:

Working backwards from understanding the magnitude of failure demand in a particular setting, failure demand must be measured in context and the impact assessed. To ensure that this measurement is valid – a clear

[1] Rather than improving capacity, either through capitalisation or raising efficiency.

framework for understanding failure demand must be constructed. To allow for this enquiry – it is important to understand the broad demand landscape in health care, in other words identify the building blocks that make up demand in health systems.

Listing these elements chronologically, to answer the research question, the following objectives must be met:

1. To create a model of the building blocks of demand as viewed systemically,
2. to create a means to identify and categorise failure demand in a complex hierarchical setting
3. using the understanding gained of how demand is structured and how failure demand responds in health care, to establish a baseline measurement of failure demand in a health system setting.

1 - 2.6. Scope and Delimitations

This work is conducted in a general health care setting, considering contexts from experience and from literature. The work is not confined to hospitals only but is to be general for understanding demand on a health care system level.

This thesis describes the impact of failure demand in a selected health care setting, in this instance in the wholesale and pharmaceutical distribution networks in a developing country. These two empirical case studies are assumed to provide an insight into previously unreported values of failure demand in pharmaceutical supply chains in developing countries.

The technical assessment of failure demand is contained to a single developing country. The context of poor supply-chain-management is assumed to be a part of the system.

Implementation, or testing implementations, of improvement initiatives is excluded from the scope of this study.

Due to restrictions imposed by the Human Research Ethics Committee which resists granting access to patients for non-clinically trained researcher, this study is purposefully scoped to exclude the need to interact directly with patients, but rather requested access to data recorded by hospital admissions systems and ERP systems for pharmaceutical supply chains.

This thesis is limited by context. Although the mental models presented are intended to be generally applicable, they are biased towards the context of familiarity of the authors. Health care systems can differ fundamentally so some of the models' categories may require modification by future researchers for their national and sectoral research contexts.

The technical assessment of failure demand is contained to a single developing country. The context of poor supply-chain-management is assumed to be a part of the system.

This study examined the systems questions related to health service delivery improvement. The authors are not clinically trained, and all clinical matters are out of scope.

Implementation, or testing implementations, of improvement initiatives is excluded from the scope of this study.

1 - 2.7. The papers

1 - 2.7.1. A note on abstracts, literature, and method

Each of the papers, which are presented in this thesis as separate chapters, contain methodological and literature sections as appropriate. Some repetition of literature sources and argumentation may exist as each chapter must ultimately stand as an independent publication; however, it was attempted to minimise repetition.

Each paper has its own abstract, and to keep the integrity of the paper, the abstracts are presented in the style of the targeted journal.

1 - 2.7.2. A note on formatting, referencing, and numbering.

Given that all papers are written for different destinations, the original formatting for all papers differs. The formatting has however been normalised for this thesis to be consistent. This includes page, image and chapter numbering throughout and reference numbering and other matters related to formatting.

Reference styles varied and thus had to be standardised to the numbered style of this thesis (one of the papers uses Harvard referencing, so considerable syntactic modification was required). To keep the integrity of the original papers, each chapter has its own references section (even though the references are numbered continuously in the text). A complete references section is presented from Page 302 onwards.

Chapter numbers are included in all paragraph labels. The chapter number precedes the hyphen in each case. Similarly, Images and tables also contain the chapter reference before a hyphen. This allows the chapters, which are independent papers to retain internal cohesion.

1 - 2.7.3. **Journals**

Paper 1: "Demand Modalities in Health Care", this paper has been submitted to The Journal of Occupational and Environmental Medicine, where it is currently under review. This is the journal that originally published the paper by Vickery & Lynch [19], the seminal paper which sets out the most comprehensive model for demand modularisation currently in . It is believed that this work meaningfully enriches their work.

Paper 2: "Understanding Failure Demand in Health Care: A Mental Model for Demand Management", this paper is targeted at the South African Journal of Industrial Engineering has been resubmitted with corrections for final review. This conceptual paper is a specialist paper on the nature of failure demand, with a specific application in health systems. Although the intention is to keep the frameworks as general as possible to ensure external validity [20] the South African context may justify this journal for publication

Paper 3: "Measuring Failure Demand in A National Pharmaceutical Supply-Chain", This has been submitted to the Journal of Service Management. This journal publishes work in applied Operations Management and particularly investigations into domains with low levels of study. The results of this paper are useful and will be of value to the scientific and medical management communities.

1 - 2.7.4. **Authorship**

All papers were written by the candidate and he is listed as primary author on all publications. Supervisors and other authors have contributed to the papers, however the intellectual and creative elements presented in these papers are primarily those of the candidate. Author names are omitted from the papers presented in this thesis.

1 - 2.7.5. **Summary**

Figure 1-1 summarises the structure of this thesis, showing the respective objectives, and how these, via sub-questions, are addressed across the chapters. This figure is reproduced larger, as Appendix 1 on Page 1. The detailed discussion of these elements is given in the preceding sections (Section 1 - 2.2 onwards).

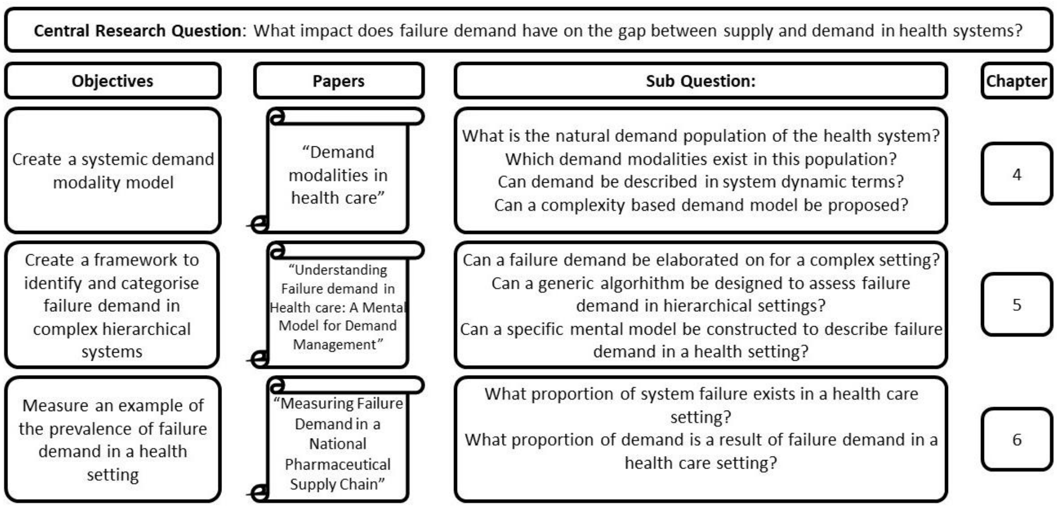


Figure 1-1: Structure of this thesis

1 - 3. References

1. Merry, M.D., *Healthcare's Need for Revolutionary Change.* Quality Progress, 2003. 36(9), p. 31-35.
2. World Health Organization, *World Health Statistics 2017*. 2017, World Health Organization
3. Worldometer. *Government and Economics Indicators*. 2020 [cited 2020 15/03/2020].
4. International Monetary Fund, *World Economic Outlook*. 2020, International Monetary Fund, Washington D.C.
5. United Nations Development Programme, *Human Development Report 2019*. 2019, United Nations, New York City - New York,
6. Unesco Institute for Statistics, *Global Education Digest*. 2020, UIS Montreal.
7. Sarnak, D.O., Squires, D., Kuzmak, G., and Bishop, S., *Paying for Prescription Drugs around the World: Why Is the Us an Outlier.* Commonwealth Fund Issue Briefs, 2017. 2017, p. 1-14.
8. Dijulio, B., Kirzinger, A., Wu, B., and Brodie, M., *Data Note: Americans' Challenges with Health Care Costs.* Kaiser Family Foundation, 2017.
9. Whitehurst, L., *Utah Sends Employees to Mexico for Lower Prescription Prices*, in *Associated Press News*. 2020, The Associated Press, New York.
10. Affordable Care Act, *Public Law 111-148*, in *Title IV, x4207, USC HR*. 2010, United States of America. p. 2010.
11. Smith, M., Saunders, R., Stuckhardt, L., and Mcginnis, J.M., *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*. 2013, National Academies Press, Washington D.C.
12. Aaron, H.J. and Schwartz, W.B., *Managed Competition: Little Cost Containment without Budget Limits.* Health Affairs, 1993. 12(1), p. 204-215.
13. Graban, M., *Lean Hospitals*. 2008, CRC Press, Boca Raton, London, New York
14. Ohno, T., *Toyota Production System: Beyond Large-Scale Production*. 1988, CRC Press, Boca Raton, London, New York
15. Liker, J.K., *The Toyota Way*. 2004, McGraw Hill, New York
16. Yin, R.K., *Case Study Research: Design and Methods*. 1987, Sage Publications, Inc., Beverley Hills

17. Levine, H.G., Gallimore, R., Weisner, T.S., and Turner, J.L., *Teaching Participant-Observation Research Methods: A Skills-Building Approach.* Anthropology & Education Quarterly, 1980. 11(1), p. 38-54.
18. Moodie, J. and Hapgood, K. *The Phd Thesis by Publication in Engineering: Insights for Supervisors*. in *23rd Annual Conference of the Australasian Association for Engineering Education 2012: Profession of Engineering Education: Advancing Teaching, Research and Careers, The*. 2012. Engineers Australia.
19. Vickery, D.M. and Lynch, W.D., *Demand Management: Enabling Patients to Use Medical Care Appropriately.* Journal of Occupational and Environmental medicine, 1995. 37(5), p. 551-557.
20. Marczyk, G., Dematteo, D., and Festinger, D., *Essentials of Research Design and Methodology*. 2005, John Wiley & Sons Inc, Hoboken, New Jersey

Chapter Two: **Literature Review**

2 - 1. Introduction to The Literature Review

The purpose of this literature review is to contextualise this study, provide background and some of the key debates that exist in the field of demand management in health systems. Figure 2-2 shows how the topics in this literature review are structured.

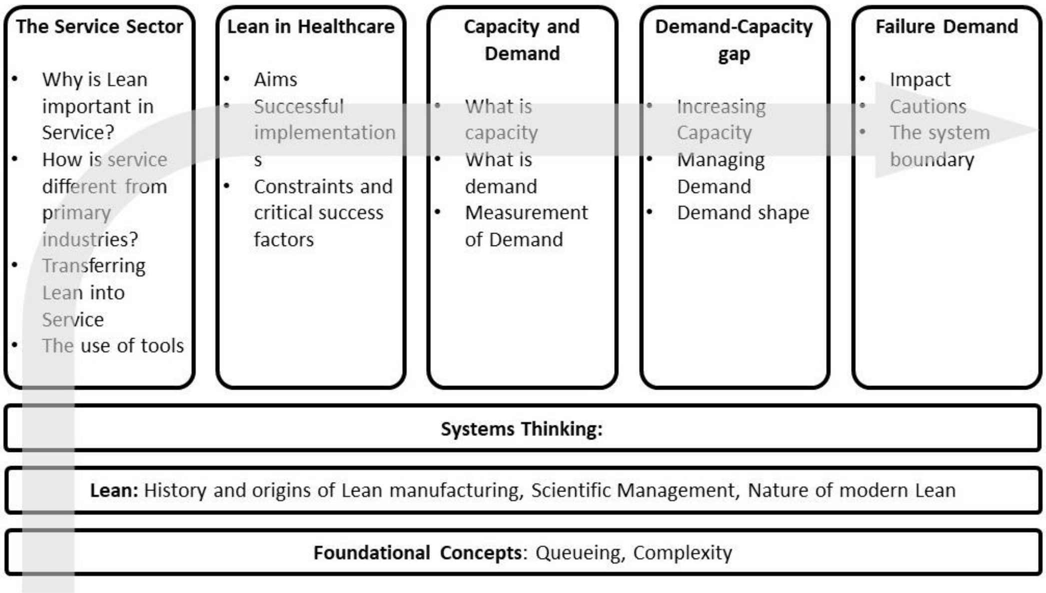


Figure 2-2: Flow of the literature section

This chapter first presents a few foundational concepts related to queueing theory and the impact of complexity. After this the more targeted review begins with an exploration of the primary theoretical lens of this work, namely Lean manufacturing, its origins and evolution. This provides an understanding of how Lean works in manufacturing. This is followed up by the complementary, secondary theoretical lens for this thesis, namely systems thinking. After that, the service sector is explored. This review explores how services differ from manufacturing and makes arguments for the use of Lean in this sector. The review then moves from the general service case, into the particular case of Lean health care. Here the use of Lean is described based on work that has been published. This section explores the common aims of Lean implementations in health care and provides commentary on whether such interventions are generally successful. Out of this arises a framework of key constraints and associated success factors for Lean initiatives.

The work to this point has moved from the general; describing Lean, then services broadly and then Lean in health care specifically. The second key part of the literature focuses on demand, capacity, queues and managing the gap between capacity and demand. Because it is a central part of this thesis, considerable attention is given to describing the current thinking around failure demand its current use in services and health care. A few cautions about the use of failure demand are raised, that look specifically at the way in which John Seddon views systems.

2 - 1.1. Literature sources

Broadly, literature was extracted from credible, peer reviewed sources. The only exception to this practice was the use of credible grey literature sources such as government documents (such as the National Health Institute from the USA), Agency and NGO documents (for example from the World Health Organization and the Lean Institute). Occasional use of textbooks and other books was used to gauge baseline and standard practices.

Literature sources were mainly found by making use of Google Scholar and PubMed, and to a lesser extent Science Direct and Scopus. Where doubt existed about the credibility of data sources, the originating journal was vetted against lists of predatory publishers [21][2] or as a positive test, against the Web of Science Master Journal List (The ISI list) [22] to ensure opinions captured in this thesis were safeguarded against questionable scientific practice.

2 - 1.2. The use of grey literature

This thesis does make use of grey literature. Given that the thesis wishes to understand the way that demand works in health systems, some of the more valuable insights into processes, behaviours and operations are captured outside of the academic press, as the incentives to publish may not exist. A great deal of this information is instead captured in documents by Governments and agencies (such as those of the Mental health Services of the Scottish NHS) or documents, policy papers and other publications of supra-governmental organisations, such as the World Bank Group, the World Health Organization, agencies of the United Nations and other reports, for example from Non-Governmental Organisations and aid agencies. Although the preference for peer reviewed academic sources remains, in their absence, and to broaden the scope of input, and to enrich the overall output, grey sources play a crucial role in the ability of this thesis to answer the questions that it sets out to understand.

[2] Jeffrey Beall – Librarian at the University of Colorado has compiled and curates a list of journals and publishers suspected of predatory practices. The credibility of Beall's list is established through the large number of good universities, including Yale and Caltech, referring their students to his list.

Grey literature is generally considered to be "the diverse and heterogeneous body of material available outside, and not subject to, traditional academic peer-review processes" [23 Pg. 433] whilst Conn et al. propose that grey Literature can be identified by "the inability to find it in computerised bibliographic retrieval systems" [24 Pg. 256]. Grey literature can exist in many forms. Adams et al. identify 35 different forms of literature ranging from unpublished academic work and census and other statistical sources, to informal communications, wiki-articles, emails and tweets [23].

To clarify the vast range of documentary sources outside the classical academic realm, Adams et al. propose a classification continuum based on the credibility of the source and the trustworthiness of the outlet [23]. In doing so – they identify three tiers of grey literature. Tier 1 represents the most credible grey sources, including government reports, books and non-academic journals [23]. Tier 2 is defined by sources of intermediate credibility and trustworthiness including annual reports, presentations and wiki articles [23]. Tier 3 sources are difficult and often impossible to check and the ability to validate claims is equally difficult; including material from blogs and emails and tweets and letters [23].

2 - 1.2.1. **Reasons for including grey literature.**

Because grey literature is heterogeneous [23] it has the potential to present greater methodological diversity, a richer spread of opinion and perspective [24]. This leads to a broadening of opinions, voices and perspectives [24]. This can add value to findings especially when used in analyses together with more traditional sources [25].

Grey literature may be good academic work that experiences difficulty in getting published, often due to the scope of and funding for a study [23, 26]. Restricted funding can lead to limited scope, which may in turn lead to a lower publication probability [27]. Despite the scope concerns, the underlying rigour and methodological suitability is often equivalent to so called "white" [23] literature [24]. This means that a great deal of worthwhile and rigorous research may be excluded from the literature that is easily accessible through computerised repositories [24, 26-28].

Even when funding and scope are sufficient, if findings are not dramatic, publication is often tricky. To publish work, journals, editors and reviewers are likely to seek statistically significant effects [23, 25, 29]. A consequence of this, is that the body of knowledge systemically excludes neutral findings or outcomes showing weaker effects [25] – which Tavola calls "Academic Misconduct" [29]. This means that the exclusion of unpublished work may distort the true picture as a full reliance on published sources only, excludes findings. This suggests that grey literature has considerable value in systematic-or meta-analyses [23, 24]. The power of grey literature in this instance is to moderate claims made in the most enthusiastic (or statistically significant) sources balancing findings [23, 29].

Perhaps of the greatest importance for this thesis is that not all organisations aspire to publish their good work in the academic literature [23, 26]. Much work is done by governments, consultants, organisations, and agencies without releasing the work for common consumption – either due to the confidential nature of the work or because sharing knowledge is not of importance for such entities. Therefore, the body of knowledge represented by entities – that have no interest in academic publishing – remains significant. This is particularly clear when trying to understand organisation-internal practices and conventions.

2 - 1.2.2. **Conclusion on grey literature**

Like Moraros [30] This thesis makes use of grey sources where published peer reviewed material does not exist. Where peer reviewed material does exist, grey sources are sometimes used to support or validate claims if necessary. Grey sources used in this thesis are generally Tier 1 documents [23], which includes government reports [23]. This aligns with the general positivist philosophical position of the research [31]. Further sources include government agencies (such as the NHS in the UK) or agencies that act on the government level, such as the World Health Organization, the World Bank[3], UNICEF and others.

A limited number of sources from Tier 2 and Tier 3 are also included in this thesis. Company internal documents are used occasionally, in particular sources from Vanguard consulting – John Seddon's company [32] which heavily focuses on failure demand [33]. This is a justified use of Tier 2 grey literature because of Vanguard Consulting's key role in failure demand – work. In one instance a personal communication with John Seddon is presented – this is an example of a Tier 3 source, however the argumentation from the conversation is also backed up by some of his published sources.

[3] The author's experience writing policy briefs for this agency revealed a thorough, rigorous and in many respects more exacting process of peer review than experienced in publication in academic journals. Equally it was interesting to discover that the majority of reports remain internal outside the public domain.

The thesis takes the position that grey literature is a viable source of insight into the operation of health care, represents a valid means of gaining an understanding of actual operating realities and allows for access to system-understanding that is not published in the traditional academic press for a variety of reasons, chiefly that the originating agencies and authors have no interest in academic publication.

2 - 2. Foundational Concepts

This section presents a few concepts that are core underlying principles that will emerge throughout this literature review and will act as the basis for commentary from time to time in the papers that make up the remainder of this thesis.

2 - 2.1. Utilisation and queues

Understanding queueing must consider the work of Sir John Kingman [34-37], whose contributions in queueing theory and system response has resulted in the simplified relationship shown in Equation 1 [38] that John Bicheno in a personal communication referred to as the "Equation of Lean" [39].

$$\mathrm{Q} = \frac{C_a^2 + C_p^2}{2} \times \frac{\rho}{(1-\rho)} \times T_p \qquad (1)$$

With C_a the coefficient of arrival variation, C_p the coefficient of process variation, ρ representing system utilisation, expressed on a scale [0…1], or load divided by capacity and T_p as the standard process time.

From Kingman's pure formulations [34-37], a change in load or in the utilisation in turn is positively associated with the queue length. Hopp & Spearman refer to this as the "law of utilization" [40]. This is more clearly shown in Figure 2-3 (adapted from [41]). This figure can be reproduced in a simple variation dice game that is described by Bicheno [42] and in a great step-by-step visual guide by Roser [43]

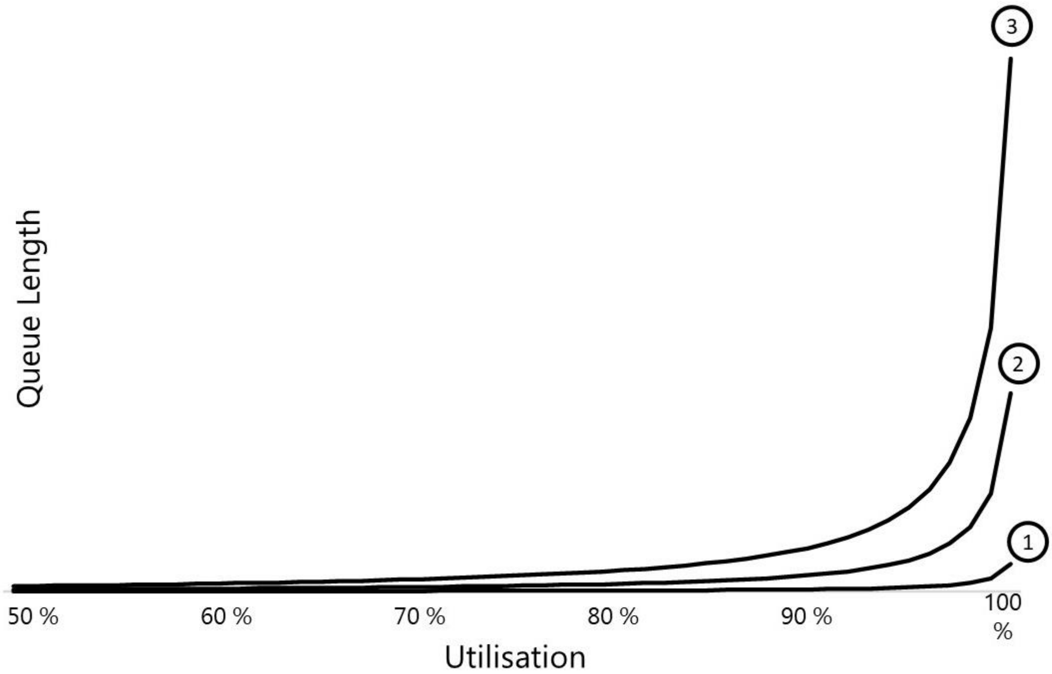


Figure 2-3: Kingman's equation - the effect of changing utilisation and variation on queue length [38, 40, 43]

Figure 2-3 shows how the queue length (shown on the vertical axis increases exponentially in higher utilisation environments. The three curves represent different variability cases. Low variability, i.e., a very predictable process in terms of demand and supply dynamics is shown by ①, a mid-line variability level is indicated by ② and high variability is shown by ③. This highlights that variation can generally be absorbed by a system at lower utilisation levels, however at higher utilisation rates, high variability leads to considerable load increases which manifest as exponentially stretched queues. Indeed, the “Law of Capacity” states that hundred percent utilization is impossible [40].

Similarly, process and arrival variation are affected by different modes of demand. This means that variation can be reduced by reducing demand. If it is accepted that demand contains elements of failure demand, then in turn this means that reducing failure demand would reduce overall queue length [39].

Variation is of concern in value chains because high variability leads to a system that is difficult to control and plan, and therefore generally leads to accumulation of demand [38, 40, 43]. Demand however is often thought of as a manifestation of an obscured black box process, however this thesis argues that the complex composition of demand is important, and possible, to understand and once demand is understood, it ca be managed [44].[32]

2 - 2.2. Little’s law and factory physics.

Little’s so called “Law”[4] [47] argues for the relationship between demand, the rate of arrival and the rate at which work can be done [48]. Little’s law is possibly the most useful relationship in scheduling and establishes a relationship between the average number of items in a system, the average arrival rate and the average waiting time[47], as shown in Equation 2

[4] Little argues that the “Law” is rather a “Tautology” [45]. Little also reveals that the most likely originating author responsible for the equation was Philip Morse [46].

$$L = \lambda \times W \quad (2)$$

With L as the average number of items in the system, λ as the average arrival rate and W as the average waiting time.

To apply Little's law to health care, a corollary [48] has L as the average length of (patient) stay, λ as the average number of patients in the process and W as the average arrival rate of patients to the system.

In Factory Physics, the relationship was rewritten with greater emphasis on production [49] and emphasises output as can be seen in Equation 3

$$TH = \frac{WIP}{CT} \quad (3)$$

With TH = throughput, WIP being the product waiting in the system and CT being the process cycle time.

The emphasis on output makes this version of Little's Law more useful for production planners. But the key realisation to allow this equivalency, is to assume that the consumption rate is equal to the arrival rate and that the WIP does not accumulate or age at different rates. As a caution, Little highlights that his law does not hold in transient periods or in unstable systems [49].

Little's law is sensitive to variation and instability [49]. All elements in the law are averages and as such conceal erratic behaviours[47]. In addition to stability, an important condition for Little's law to hold – is conservation of flow [49], or what Hopp & Spearman refer to as the "law of conservation of material" [40]. A simplified way to understand this, is the idea that what goes in, is what comes out. A caveat from Factory Physics is a provision for yield losses [40] which can be thought of as friction in physical systems. Figure 2-4 shows a hybrid (developed for this thesis) of Little's law and the law of conservation of material [40, 45, 47].

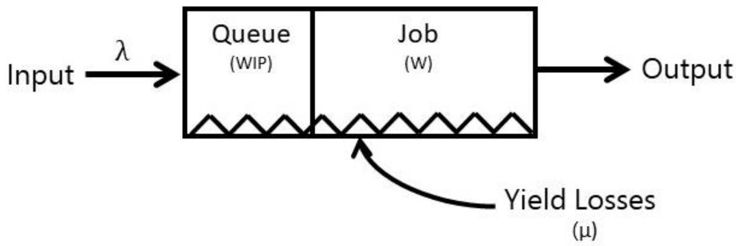


Figure 2-4: Hybrid model showing the views of Little, Hopp & Spearman [40, 47]

Figure 2-4 shows that any process has an incoming demand, arriving at a certain rate (λ), enters the Work in Progress (WIP) queue and finally exits the system as completed product. The yield losses (shown as a "rough" surface (μ)) are Hopp & Spearman's difference between input and output and the reason why inputs will always be greater than outputs.

Notice that for these systems, and also in the case of Kingman's more elaborate models (see section 2 - 2.1) beyond yield losses, nothing exits the system unprocessed [40, 45]. This is an important gap, as it does not reflect reality. In manufacturing systems, rework and repair churn add to WIP levels [15, 50]. Similarly, in service systems failure demand [44] becomes an important path for work to leave the system incomplete and to re-enter again later as demand, also a churn – model.

2 - 2.3. Complexity and the Cynefin framework

The Cynefin[5] framework was designed as a tool to understand and model the level of complexity and its impact [52, 53]. Dave Snowden makes recommendations to leaders, how to respond in complex situations in his Harvard Business Review article [54]. In recent times, Cynefin has been used to solve problems in numerous industries, ranging from Information Systems [55] to creating dynamic management policing strategies for riots [56] and assisting a pharmaceutical company to create new product strategies [54]

The Cynefin framework is generally drawn up with curvy lines [57], perhaps, to imply that the boundaries between categories are fuzzy. Unlike most 2X2 matrices [58] (of which the most famous is probably the so-called Boston Consulting Group (BCG) – Matrix [59]) the Cynefin framework does not have a Cartesian intersection point at (0:0), but rather has an indistinct region in which the user concedes themselves defeated to try to categorise a situation into one of the regions. This region is called the area of "disorder" and is comparatively poorly described in literature.

[5] Welsh for "Habitat" or "Place" and is pronounced Kooh – neh - fin [51]

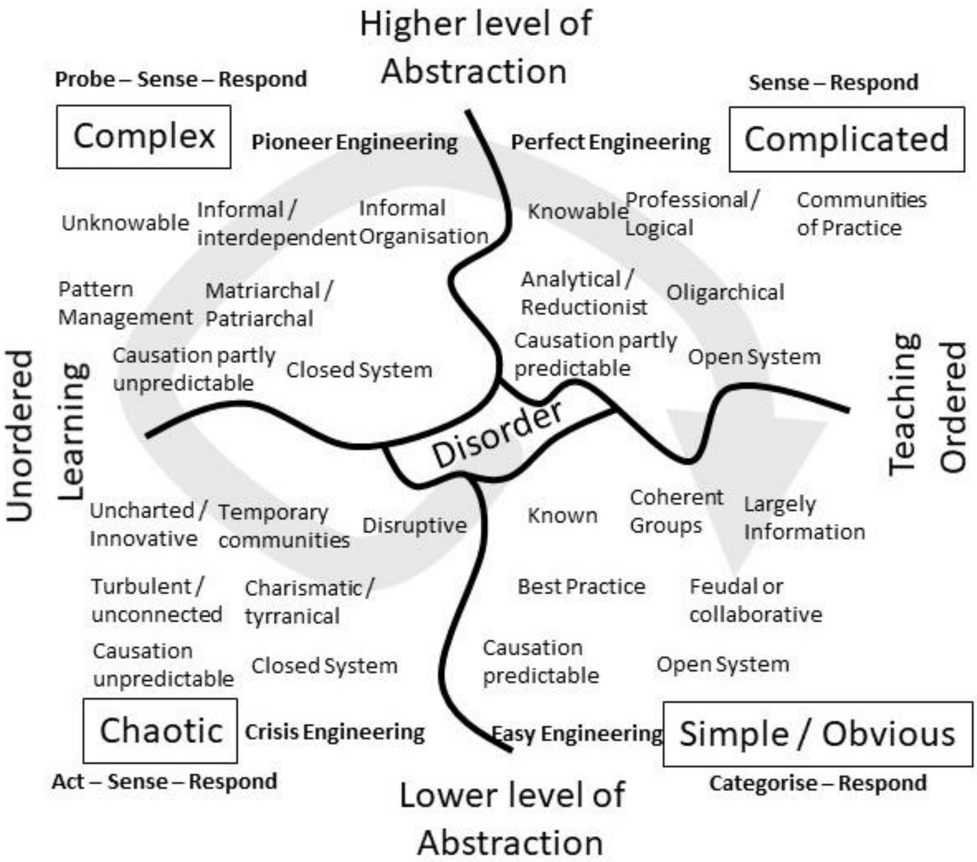


Figure 2-5: A synthesis of sources describing the Cynefin model [51, 53-55, 57, 60, 61].

Figure 2-5 is our synthesis of multiple sources, which includes numerous references to work by the creator of the Cynefin framework; Dave Snowden with a particular emphasis on "complex acts of knowing" [57] and supplemented by his numerous other publications [52-54, 62]. These sources create the general structure, the interpretation and meaning for the regions and is augmented by other work, most notable Siemens [60] and further reinforced by Brougham [51] and Hasan[55, 61].

2 - 3. Lean

Although this is not explicitly a study in the domain of Lean, the theoretical lens that has shaped the thinking of the author is explicitly a Lean one. Lean prioritises value from the point of view of the customer [63], it achieves this, through five key "pillars" [64], namely creating value, then designing value streams and the flows that enable these. To achieve the flow requires that the customer pulls production and only then can perfection – a key philosophical aim of Lean, be achieved [14, 15, 63, 65]. It is from the point of view of this philosophy that the author's lens has been shaped. The following sections highlight some of the underlying elements of the Lean philosophy, and a tacit justification for the research presented here.

2 - 3.1. History of Lean manufacturing.

Lean, as a term, likely originated in the corridors and common room of the Sloan School of Management at the Massachusetts Institute of Technology (MIT). The first use of the word, with its current meaning, was in a Masters dissertation in 1988 by Krafcik [66]. Three years later, his colleagues at MIT (Womack & Jones) published "The machine that changed the world" - the book that is generally credited with the introduction of the term, and certainly made it famous [65].

To fully understand the history of Lean, however, one must separate name from action and from the underlying philosophy. It is generally agreed that the purpose of Lean is to put the needs of the customers first [63, 67]. This philosophical position needs practical approaches to achieve this purpose. And many of these practical approaches have existed throughout history.

Seeking the earliest signs of what is now thought of as "Lean" thinking is an interesting act of selective interpretation of industrial history. Several authors have identified behaviours that resemble elements of the modern Lean philosophy. This includes practices from antiquity to the industrial revolution with a few waypoints in between. Although the practices may resemble Lean, the collective whole of the discipline took far more time in its construction.

These practices include the first mass-production shipyard – the Venetian Arsenal [67] which was the largest mass manufacturing enterprise in Europe prior to the industrial revolution which gave the Venetian kingdom dominion over the Mediterranean as a military- as well as a merchant-power [68].

Later examples of sophisticated industrialisation include the work during the industrial revolution of the father of modern economics, Adam Smith, who broke down complex tasks previously performed by artisans into simpler, replicable tasks, which improved quality, throughput and reduced the level of skill required to produce [69]. But Smith's contribution was more than that – as he stated in "the wealth of nations" that:

> *"Consumption is the sole end and purpose of all production; and the interest of the producer ought to be attended to, only so far as it may be necessary for promoting that of the consumer" [70]*

Herein Smith illuminates his unknowing alignment with the philosophy of the primacy of the needs of the customer, and even suggests the idea of production being pulled, rather than pushed. Ideas that would later drive Lean thinking.

It was this thinking, whether consciously or not, that drove Whitney [71] and Whitworth [72] towards standardising interchangeable parts.

In the early nineteenth century, in the USA, Whitney patented the cotton gin, and the use of interchangeable parts, allowing for simplified mass production. The mental leap towards interchangeability was possibly one of the most definitive transitions that allowed the immense prosperity to emerge in the southern states of the USA [71]. Later Whitney used the same principles of interchangeability to deliver high volumes of low-cost weapons to the US government [73].

In the mid nineteenth century, in the United Kingdom, Whitworth extended interchangeability with standardisation – by setting norms for nut- and bolt-threads [72]. Before Whitworth (whose name remains as part of the BSW (British Standard Whitworth) all threads were custom cut by artisans for a one-to-one fit. Limiting interchangeability, quality and work-pace.

2 - 3.1.1. **Scientific management – the seed of industrial engineering and of Lean**

Towards the end of the nineteenth century, it was thought that management should essentially be military [74]. The role of the manager was to be filled by a technical expert whose administrative skill was of secondary importance. This expert would have the twofold role of heavy handed control, which could approximate bullying, and on the other hand, rested heavily on the gut feel of the manager [75]. At the time, management was thought of as something that could not be taught [76]. Gilbreth laments that "an idle workman is seen as a disgrace" [74 Pg.29], saying that to prevent this, exhausted labourers would rest whilst they worked whilst Fayol found it extraordinary that managers would go as far as to monitor their workers' church attendance [75].

This is the context into which the three pioneers of so-called scientific management philosophies entered. From modest roots, Frank Gilbreth, later joined by his wife Lillian published a primer on scientific management [77], a year after the "Principles of Scientific Management" [78] was published by Frederick Winslow Taylor.

Gilbreth began his career as a builder [79] and realised the efficiencies that could be gained in brick laying by modifying scaffolding heights [79]. Whilst Taylor started his career as a pattern maker and machinist, thus gaining an understanding of how a factory works from the inside.

Although the principal pioneer of the scientific methods is generally recognised as Taylor, the work of Frank and Lillian Gilbreth and Taylor is almost identical, in the time of publication as well as in the content and impact. Both sides focusing on fatigue [78, 80, 81] and work study [74, 78, 81, 82], standards, standardisation, work design, compensation, rest and efficiency [ibid]. Indeed, the debate of who is the father (and mother) of scientific management, is perhaps as energetic as the disagreement over whether Newton or Leibnitz are the father of Calculus [83, 84]. Much debate exists today and mimics the snide personal conflict that existed between Frank and Lillian Gilbreth and Frederick Taylor.

The debate notwithstanding, it is worthwhile to foreground Gilbreth and Gilbreth's "Fatigue Study, the Elimination of Humanity's Greatest

Unnecessary Waste: A First Step in Motion Study" which refers to the loss of productivity through overload, and unevenness as "Humanity's greatest Unnecessary Waste" [80]. This book introduced a great amount of the Lean language, including the idea of waste (Muda), that waste can be necessary or *unnecessary* and indeed that fatigue, a consequence of overload (Muri) and unevenness (Mura) leads to the waste of human potential, a waste that was only re-recognised in the Lean literature [64] much later than the famous seven wastes [14].

Led by the scientific management approaches introduced by Frank and Lillian Gilbreth and Taylor, Henri Fayol [76] was a pioneer who started to think about organisational structure. Fayol looked at management processes– an evolution of the procedural approaches of the scientific managers. Fayol proposed that organisations, and people, should be led with foresight – which would in modern language translate as "strategic thinking".

Perhaps the key contribution by Fayol, was his five tasks that constituted a manager's job [76]. These roles include Planning and Organizing [75], but more importantly for this thesis, Fayol emphasised "Command, Coordination and Control" as crucial management skills [75]. This is an essential transition, as the base thought of failure demand is contained in John Seddon's 2003 book – "*Freedom* from Command and Control".

This kind of thinking was simultaneously being developed by Max Weber, who advocated bureaucracy. The word at the time didn't attach the modern pejorative associations, but was rather associated with the merit based assignment of function on rational principles, rather than polito-hegemonic decree [75]. In a way – the idea of bureaucracy at the time implied that your role was determined by how good you were rather than by who you were. This expressed the idea that the best man (and at that time, it really, always was a man) for a job, is the best man for a job. This thinking is expanded on by Taylor, who again with terrific vision of future Lean thinking, states that "In the past the man has been first; in the future the system must be first" [85 Pg.7].

By the early twentieth century, Henry Ford was able to use the mass production thinking from the middle ages, the industrialised use of man and machinery, interchangeable and standardised parts from the nineteenth century and the contemporary scientific management approaches, and the nascent thinking around systems, to create the world's first moving production line [86]. Many of the features of Ford's production system are reminiscent of modern Lean, including the elimination of waste, a passionate frugality that mirrors modern thinking around efficiency, just in time delivery of goods and other tools [87].

The Toyota Production System (TPS) [14] arose in response to the resource and machinery constraints that were brought about by the Second World War; the destruction of manufacturing capacity and facilities and the loss of human resources [15].

TPS was pioneered by Taiichi Ohno who was joined by American manufacturing guru, J Edwards Deming. Ohno and Deming created the structured formulation for TPS, which prioritised to put the customer first, to sustainably cut costs, improve production and quality whilst fostering a culture of employee involvement [15, 63].

The philosophy underpinning the work of Ohno and Deming was able to lift a shattered Japanese economy out of ruin by making the Japanese manufacturing industry efficient and world class. So much so, that Womack & Jones, when introducing the term "Lean", singly attribute Japan's modern day economic prosperity to the emergence of these innovative efficient practices in post-war automotive manufacture [65].

2 - 3.2. Nature of modern Lean

Lean as a philosophy attempts to make do with available resources. This is achieved by enlisting the totality of the staff of the organisation in the wellbeing of the company. When successfully implemented, such an approach leads to continuous improvement, the elimination of waste[6], lower cost, higher efficiency and lower defect rates [15] .

Hopp & Spearman claim that Japanese culture is naturally predisposed to being successful at Lean – given the almost stoic, and certainly frugal character required to deal with resource and space constraints [40]. Although many claim that Lean only works in Japan, this is a misconception [88], and numerous examples of exemplary Leanness exists globally [88].

It was found that South African "Ubuntu" culture shares many of the frugal collaborative elements of Japanese culture – leading to an underlying ability to thrive in deprivation as a team [89].

Although Lean consists of many "tools" [63], which are often taught as the essence of Lean, the true core of Lean is about the philosophy that drives a

[6] The key wastes identified in Lean literature are: Overproduction, Over Processing, Transportation, Motion, Inventory, Waiting and Defects [63]

culture of people involvement, leading to continuous improvement and reducing waste to elevate the customer as the principal driver (puller) of the process [14].

Lean focuses on the efficient use of resources, material, machinery and people. The classical counterpoint to this is an organisation that hides poor planning and an ineffective system by having an excess of material, machinery and people. The Lean philosophy achieves success by emphasising the needs of the customer and modifying an organisation to meet these needs [63]. This is achieved by focusing on the five key pillars; recapping: creating value from the point of view of the customer, creating value streams; the means of delivering the value to the customer, creating flow in the value stream, to ensure efficient delivery, creating pull, thus reducing unnecessary accumulation and creating a responsive system, and achieving perfection [90].

Womack & Jones identified health care as one of the disciplines which would benefit from the Lean principles [90]. Following on from this, several key texts emerged in the field of Lean in health care. Mark Graban's Lean Hospitals [13] introduces Lean to health practitioners, but is weak on implementation, a gap which is addressed by Zidel [91] in his practitioner's guide.

This thesis frames the philosophies and the work as a hybrid that primarily draws on theory from Lean and makes use of Systems Thinking as a supporting theoretical lens to investigate health system. The next section takes an overview of the literature on systems in general, which gives rise to systems thinking.

2 - 4. Systems

The pace of publication in the field of systems thinking has accelerated considerable over the last three decades. In a systematic review that analysed the period 1990 – 2015, Williams et al. found 96 papers in the discipline, however almost all papers were published after 2000. In a larger study by Rusoja et al. covering the period 2002 – 2015, 516 papers were found, starting with ten papers in 2002 and finding more than eighty papers published in 2014. The growth in recent years has been called "exponential" [92 Pg. 878]. Major categories that were identified from these publications were dynamic systems, emergence, complex adaptive systems, and interdependence. Being outside the scope of this thesis, they will not be discussed further save to note that they have common characteristics that fall within the central elements of systems thinking theory whose features are mentioned below.

2 - 4.1. Central elements of systems theory

The early work on Holism by Smuts [93] proposes the central premise that understanding phenomena requires broad complex thinking, rather than reductionist simplification. Out of this emerges the broad field of systems theory [94-96], of which systems thinking is a key constituent [97].

Systems theory responds to unexpected and counterintuitive phenomena where linear logic cannot be used to understand behaviours. For example in a wetland less water evaporates by transpiration than in individual plants [97] this may be because the surrounding plants' transpiration raises the environmental humidity, thus reducing this process [97]. Similarly, Oxygen and Hydrogen behave completely differently to when they are joined together in water [98]. This becomes even more interesting when ice, liquid water, and steam all behave differently due to environmental conditions. Oxygen and Hydrogen, seen in isolation of their relationship to each other and the environment within which they exist can define the difference between a cube of ice and a vigorous explosion!

These analogies describe some of the important features of systems, and why being able to apply a systems-lens to complex problems is important. Systems thinking emphasises that the interconnectedness of elements is central to the holistic functioning of an entity [96, 99-101].

2 - 4.2. Systems thinking

Although systems theory and systems thinking are strictly distinct fields, in the academic literature related to organisations, the distinction is unclear and the two are almost entirely conflated. This section continues with the use of the terminology systems thinking from here on.

2 - 4.2.1. Interconnectedness and interdependence

The first principle of systems thinking is that systems do not presume independence of agents, there is interconnectedness [1, 11] and interdependence [11]. Systems are about the collaborative nature of elements. In the same way that water does not behave like elemental oxygen and hydrogen [98], so a system of clinicians, nurses, physical infrastructure, money, information and a multitude of other actors does not behave like the individual elements. Cause and effect relationships exist between entities inside a system, but describing this in a linear fashion is not always possible [102]

2 - 4.2.2. **Dynamics, non-linearity and feedback loops**

A microphone held in front of a speaker creates a feedback signal. The microphone 'hears itself', amplifies this signal and repeats this cycle many times. What starts as inaudible ambient sound rapidly becomes a soft hum and quickly turns into an unbearable screeching sound [103]. Removing the microphone from the speaker however rapidly damps the system back down to silence. This is a familiar example of how a system, feeding back onto itself, can completely alter the system response and behaviour.

A key characteristic of systems is that the current operation of the system can impact the future operation of the system. A practical example from this review, is the rising utilisation curve presented in Figure 2-3. As with a microphone, incorrect actions taken will raise utilisation, which will disproportionately raise the queue length. This in turn creates an entirely different operating reality in the future. This can continue to escalate uncontrollably over time. In an unconstrained system, such a reinforcing trend can continue essentially undamped [92, 104].

There are two types of feedback, one type is a reinforcing signal, that strengthens a system response [99, 104]. In the microphone example, both the rise in noise as well as its abatement, are reinforcing signals. Whatever behaviour exists in the system is strengthened (or weakened) [99, 104]. The other type of feedback is a balancing feedback, which is negatively associated with the signal. In the microphone example, if the amplifier has circuitry to limit feedback by not allowing it to go above a certain level, this takes the signal and filters it to an acceptable level, and keeps it at that level [103]

The ideas around feedback [92] and dynamics [105] emphasise that systems are impermanent and that responses and behaviours can change with time and that frequently systems are responsible for their own dynamics.

The idea of non-linearity [105] is crucial to understanding systems. Lorenz' "butterfly effect" [106] is a chaos-theory thought-experiment that proposes that a butterfly flapping its wings can via multiple cascading events cause a hurricane or alter its path weeks later. It is perhaps wise to not consider this effect to be a literal event, but rather an illustration of how signals can have a non-linear impact on system behaviour. The amplification of a stimulus is reported in literature sources [105, 107] and in other examples in this section.

Given that systems experience dynamic interactions continuously from feedback and external and internal stimuli, it follows that successful systems are by nature prone to being able to respond quickly to change, to adapt and to be resilient.

2 - 4.2.3. **Self- organising**

Self-Organisation refers to the idea that systems undergo change due to the stimuli they experience [92]. Out of chaos, order is established, for example crystallisation from a disordered solution into a regularly structured crystal [108]. Luhman [109] shows how systems spontaneously tend towards some sort of order, such as for example traffic circles, which create order instead of the anticipated chaos.

Because systems operate as complex interactive entities, with resilient responses to feedback and dynamic behaviours, systems change their characteristics as a response to external and internal stimuli [97]. These together incrementally build behaviours and responses to stimuli separated from formal command and control structures [44].

2 - 4.2.4. Emergence

Emergence is an important characteristic of modern systems theory [97]. Emergence is the notion that a system becomes more than the sum of its parts, in other words, a system is able to acquire competences and characteristics over and above the elements that constitute it [97].

2 - 4.3. Synthesis

Systems thinking is a lens in which the totality of a phenomenon must be considered [92]. This includes the fact that a system is a complex network of interconnected parts [105]. These parts create order autonomously to allow for effective operation. Out of this self-organised whole come the system's characteristics that exceed the sum of its members [97]. These emergent characteristics are the 'person' of the system.

Using the Lean and systems views as a theoretical lens, literature presented in the next sections explores the service industry, which is a complex system with considerable difference to traditional manufacturing contexts. The service context is refined by exploring the health care industry specifically.

2 - 5. The Service Sector

Along its genesis, as was discussed, Lean has generally been a strategy for the manufacturing industry, however recent developments have also placed Lean into the service industry. This section deals with the placement of Lean in services broadly and then focus on Lean health care specifically.

Piercy & Rich stated in 2009 that Lean work in the service environment is "rare" [110]. This sentiment was echoed by Gupta et al. in 2016 [111] when they referred to Lean in services as "making [in]roads, though the research [was] still at the nascent stage" [111 Pg. 1025]. These "inroads" can be seen in the fact that published academic literature on Lean in Services accelerated rapidly from approximately 2010 onward [111]. In a systematic review, published in 2016, Gupta et al. analysed the landscape of Lean in the service sector broadly [111]. They concluded that Lean in services can be highly successful and that a wide cross section of published results support this [111].

Several theorists have highlighted how services are different to primary and secondary activities [112-116]. Modern services tend to be "knowledge work intensive" [111 Pg. 1045], are generally for people by people [117] and the products can rarely be stored [112]. As a consequence a great deal of the success of any implementation lies in human factors [69, 118].

Due to the differences between services and other industries, Gupta et al. echo the caution of Radnor & Walley [115] that implementations cannot be done in the same way as in manufacture, where Lean originated, and that particularly, the selection of tools has to be nuanced and sensitive to the context. Put another way, it is not about using a tool for the sake of using the tool, but rather, using a tool for the sake of the job that must be done [44, 111]. This emphasises the importance of building implementations based on a deep understanding of the philosophical underpinnings of Lean, rather than the unthinking use of tools, which Seddon disparagingly calls the approach of "tool heads" [44, 119]. Although this is presented here as an admonition for practitioners in services specifically, the underlying principle: that the philosophy is prime, applies across the spectrum of Lean implementations – from the traditional manufacturing sector into service and other domains [14, 63].

The philosophical basis for Lean (in manufacturing) was captured by Womack, Jones & Roos [65] and maintains the primacy of the customer as proposed by Ohno [14], the notion of value, and the elimination of waste and a focus on the process [120], all with the philosophical grounding of the five so-called principles of Lean [65].

Recognising the idiosyncrasies of services, Womack & Jones [121] reframed their five principles of Lean [65, 90] into six different principles that apply more specifically in the service industry and call this "Lean Consumption" (which can be read in Section 2 - 5.4.2). They stretch the value chain of organisations further to the final customer-facing "service" function of the organisation [121, 122]. This echoes Levitt who proposed that all organisations are ultimately service organisation [113, 114] and supports Allway & Corbett's method of designing end to end solutions [123].

2 - 5.1. What is the service industry?

The Oxford English Dictionary defines the "service industry" as "a business that does work for a customer, and occasionally provides goods, but is not involved in manufacturing" [124]. Levitt in his influential paper on bringing production line approaches to service, proposes that all organisations are service organisations [113]. Levitt uses the example of General Motors, which, though a classical production company, probably relies more on its service infrastructure (dealerships, service centres, sales agents, marketing, finance and others) to make money than the actual production of vehicles [113]. Given this – it is perhaps wise to reframe the Oxford Dictionary definition to substitute the word "business" which implies systemic categorisation, with "function" which shows that in modern businesses, service and other functions generally coexist within single organisations (or businesses), because, as Levitt says "Everyone is in service" [113 Pg. 42].

Services have been categorised into functional domains, as shown by Cobra ([125] as translated by Leite et al. [126]). These categories are:

- "business services", which could include accounting,
- "commercial services" such as retail,
- "Infrastructure services", such as telecommunications,
- "social and personal services" such as restaurants and finally
- "public services", such as education and health.

In their systematic work, Gupta et al. show examples of Lean in each of these domains, yet find that most of the published work in Lean service systems focuses on health care [111].

For this thesis it is useful to realise that in Seddon's view, services are not about providing physical products [44] but rather that the emphasis is on what Piercy & Rich call the "pure service environment" [110]. This reveals an important, but tacit, difference in the thinking by Gupta et al. [111], Womack & Jones [121], Levitt [113, 114] and others who hold that service can have a physical product, even if it is less substantial, such as fast food [111, 122].

2 - 5.2. Why is Lean service important?

Already in Greek times, services emerged as major economic activity, with the introduction of formal education [126]. The service industry carried the greatest part of the economic activity for centuries until the industrial revolution, aided by machines, enabled the grand scale emergence of primary and secondary industries [126].

Services however later re-emerged – and approximately since the turn of the previous century, services have been the dominant, and most rapidly growing economic sector in developed countries [114]. George states that up to 80 % of the US GDP is derived from the service sector [127], whilst Leite et al. show that the overwhelming portion of global GDP now derives from services [126]. Despite the importance of the service sector, until the 1980s, the manufacturing industry remained the industry that had the greatest energy and scrutiny applied to it for improvement [128]. Allway & Corbett identify the emerging pressure on service organisations. They highlight rising customer expectations, revenue pressures, competition, increasing expense and regulatory pressures as drivers for the need to find solutions in service that mimic the undebatable gains [129] of Lean in pure manufacture [123].

Despite the scale of services, and the contribution that they make to the global economy, Levitt laments that "so called" (sic) services have often been measured by pre-industrial standards of value, derived from the sense and experience they provide [114] and that they have a natural tendency to be wasteful in the interest of ensuring experience. He elsewhere identifies that services tend to be "done by individuals for individuals, usually on a one to one basis" [113 Pg. 42] a formulation supported by Bowen & Youngdahl [117]. Levitt highlights strong examples of how hard-, soft- and hybrid-technologies enable the industrialisation of services, with all the benefits of scale and efficiency as can be referenced against in the industrial world [114]. In particular, the soft and hybrid technologies are where Levitt introduces some of the ideas that would form part of the Lean philosophy, and how their inclusion leads to the improved capability of organisations to deliver services [114].

2 - 5.3. How is service different from primary and secondary activity?

In his influential article, Sasser attempted to understand how demand works in a service organisation. Sasser's key observation centred around the different nature of demand compared to primary and secondary industries. He identified four key features that differentiated services from other industries [112].

These are:

- Services are direct – in other words, they are consumed at approximately the time when they are produced, and cannot be stored, leaving organisations with a reduced opportunity to prepare for demand.

- A high degree of producer-consumer interaction, in other words, the completion of the service is impacted by the simultaneous actions of producer and consumer of the service. This may raise uncertainty, but also make standardisation much more difficult.
- The consumer must move to where the service is provided, as services cannot (generally) be transported.
- Intangibility of what a service is, causes difficulty to measure demand and assess what actual capacity is.

Although Sasser's work remains frequently cited in literature, all four characteristics of services are not always mentioned, and some are conflated. The most frequently emphasised is the notion of the impermanence of a service. The idea that a service cannot be produced in bulk, stored in a warehouse and released upon customer demand [44, 122]. That the "product" cannot be moved to wherever a customer may be, or as Roy et al. say, a service "cannot be owned, stolen or returned" [130 Pg. 13]. Additionally the product is defined not by a "nominal limit" as set in manufacturing [131] but rather the nominal limit for the product is set by the interaction of the customer in the process [44]. What this means, is that the customer is instrumental in setting whether a service is satisfactory [44] but also that this may not necessarily be what the service provider believes the service should be [121, 122]. This means that the service "product" is far less rigid and therefore less able to be standardised than in manufacturing [44]. Or as Teece says, people do not want just a product, but a solution to their perceived needs [132].

Most manufacturing has a service function downstream in the value chain, and this service element allows the manufactured goods to make money [113, 114, 122]. The literature generally fails to mention that services are seldom on the Business to Business (B2B) level [133], whilst manufacturing with few exceptions generally is. This is a further element of the service sector that could be explored as a further fundamental difference and as a tacit driver for the importance of the customer's voice as a co-creator of the notion of "value" [122, 134].

2 - 5.4. Transferring Lean from manufacturing to service.

Bowen & Youngdahl are not sure that Lean translates into the service environment at all, however they make strong arguments to show that with sufficient discretion in selecting an approach, it could [117]. Subsequent work shows that a careful nuanced implementation can work. Gupta et al. say that "the translation cannot be mindless or literal" [111], Which resonates with Spear [118] who says that any transfer into the service industry can only be successful if both the customer and the nature of services is considered. Seddon's body of work is evidence that Lean principles and philosophies, can be successfully and sustainably translated into a service environment.

Gupta et al. refer to the years before 1998 as the "pre-lean" era [111 Pg. 1031]. This is before the thinking from the "Machine that Changed the World" [65] became mainstream, however scientific management and other precursors to Lean, (such as TPS, TQM, JIT and others) were already being practiced. In 1969, Skinner asserted that the manufacturing industry could serve as the prototype for developmental interventions in the service sector, targeting inefficiencies and designing interventions for best practice [135].

Leading on from Skinner [135], during the "pre-lean" 1970s, several important papers on the service industry were published in the Harvard Business review. Notably, two paper by Levitt that dealt with the idea that service could, and should, be "industrialised" [113, 114] and Sasser's paper that unravelled the idiosyncrasies of demand in service [112]. Although these papers all preceded the formal introduction of "Lean" they recognised the underlying early work in TQM and other predecessors of Lean. They emphasised a focus on quality and the use of tools [113] as a means to achieve greater performance in an industrialised service environment [114] that is better able to service the needs of customers [113, 114].

The next section shows approaches for carrying out Lean initiatives in a health care setting.

2 - 5.4.1. **Traditional Lean vs service Lean**

It is an almost convenient coincidence that Womack, Jones & Roos' famous book is often colloquially referred to as "The Machine" [65] . The work presented in this book focused on the automotive industry, and although it introduced the idea of Lean to the world at large, it presented it quite tightly contextualised inside the manufacturing industry (and Toyota specifically), a space dominated by machines. Ohno's ideas [14] fit comfortably into the manufacturing framing, as his own background also lay there. Although the authors think of the machine as the *system* out of which Lean cannot be split, it is true that "the Machine" is full of physical machines. As has been seen though, the thinking of "the Machine" may need some refinement to play a more direct role in the services industry.

In a Harvard Business Review paper, Womack & Jones speak of their frustrations in engaging with a service system [121]. Their language in the introduction comes across as irritated, which is precisely the user experience in many service interactions. They use words like "frustrating", "refuses", "countless loops", "miscommunication", "travel", "waiting", "without", "wanted", "tiresome", "neither help nor support", "time", "hassle" [121 Pg. 59]. They use this sense of frustration to show that Lean manufacturing does not work in its pure form in services.

2 - 5.4.2. **Service Lean and consumption Lean**

To build a bridge into services, they introduce the six principles of "Lean Consumption", which are to be seen as a "complement" to their principles of Lean manufacturing. The reformulated principles for service Lean are [121 Pg. 61]:

- Solve the customer's problem completely by insuring (sic) that all the goods and services work, and work together,
- Do not waste the customer's time,
- Provide exactly what the customer wants,
- Provide what is wanted, exactly where it is wanted,
- Provide what is wanted, exactly where it is wanted, exactly when it is wanted,
- Continually aggregate solutions to reduce the customer's time and hassle.

What emerges strongly from these principles and the remainder of the article, is that the role of the customer in determining the service that they seek is crucial. Which Seddon expresses by saying that "in service organisations, it is the customer who sets the nominal value" [44 Pg. 17].

Womack & Jones describe a scenario of a service provider providing the same task over and over (yet efficiently) dealing with the same issue. This formulation has similarity with Seddon's earlier notion of "demand that we do not want" [136] or later "failure demand" [44, 119]. Womack & Jones emphasise that solutions should be found by the right person (who should be equipped to solve these problems) and that this should dig down to the systemic root, which is essentially the argument made by Seddon [44].

2 - 5.4.3. **The Vanguard method**

Seddon shows the key importance of Lean in services, but emphasises that this should not be a selection of tools, but target improvement from the broad philosophy of seeking the "nominal value" [44], as determined by the customer. This can be achieved by making use of the Vanguard-method, the technique named for John Seddon's consulting firm. Broadly, the steps in the technique are called "Check" – "Plan" – "Do" [32].

The "Check"-phase focuses on finding knowledge [32]. Elsewhere, Seddon talks about this being the step where customer demand is studied, clarity is sought on what "value demand" is, and what the incidence of "failure demand" is [44]. It is explored whether demand is predictable [44]. The idea of predictability is essential to understanding the Vanguard method [32]. According to Seddon, if demand has a predictable it must be managed, whilst if it is not – it should be ignored. Once this has been understood, together with an understanding of current processes, the method moves onto the next phase.

The "Plan"-phase is the design phase [32]. This is the phase in which current demand patterns are integrated into a redesigned service, based on the customer's view of "value" [44]. Emerging out of this is then system in which the dysfunctional command and control elements are deleted and replaced by that what is required to run the work from a systems point of view [32].

In the "Do" phase, the work is done, plans are implemented and changes are made [32]. This is traditionally the phase to which organisations jump with too great haste, and is broadly the domain of those who favour the use of tools over the principles and the philosophy [44, 65, 119, 122]. On the presumption that the intervention has been well designed to respond to a well-studied and understood system, the intervention should be able to make systemic corrections to the organisation, thus improving overall performance [137].

2 - 5.5. The use of tools.

Seddon expresses his "antipathy" [119 Pg. 1] towards the use of tools as the means by which results are achieved. He emphasises the primacy of perspective and philosophy – the underlying understanding, on a fundamental level, of how to achieve the general goals of a more responsive, able, efficient and profitable business [14, 15, 65, 119]. He claims that if the understanding, perhaps it can be called "dogma", is taught, people will "'beat a path' to the [tool] cupboard door". This is of course then only true if the tool is likely to be helpful in achieving the goals [119 Pg. 1], whilst tools used by rote are weak [44, 119, 137, 138].

Gupta et al. concur, saying that “Attempting to go Lean by just applying a Lean toolbox of techniques in a silo fashion does not work for long-term gain. It becomes very important for a Lean practitioner to focus on the complete business value chain, rather than having a narrow ‘project-only’ focus.” [111 Pg. 1034]. Seddon makes the point even more strongly, when he says that “the codification of methods as tools obviates the first requirement to understand the problem and more importantly, to understand the problem from a systems perspective” [119 Pg. 15]. The opposite: approaching improvement from the point of view of tools first, rather than dogma first, is a common approach, but is a major source of the failure of improvement initiatives [119, 139].

Gupta et al. found that the use of tools used was very similar to those used in manufacturing [111] however they reiterated the caution that tools used by rote have little chance of success [44]. They identified many tools that had been used. This list includes value stream mapping, eliminating waste, standardisation, visual management/visual control, 5S, HR management and Kaizen. The important finding however was not the specific tools that appear to be frequently used in service interventions, but rather that there is a breadth of tools used which when suitably picked, transfer efficiently from manufacture to services. It is interesting to contrast this to Seddon’s argumentation that 5S and VSM in particular are almost purposeless acts in pure services [119]. The main reason for Seddon’s discomfort with 5S in services is that its purpose in manufacturing is to standardise the work environment to provide a levelled environment within which to make actual improvements [63]. Seddon however claims that applying standardisation to a service environment is in many ways futile, as the complexity and variety introduced by the nature of services does not actually permit standardised operations [119]. Conversely, forcing standardisation harms a system's ability to absorb variation. Seddon claims that standardisation as the genesis of improvement may lead to services deteriorating, an increase in costs and harm done to morale [119].

2 - 6. Lean Health Care

This section presents a sweep of the landscape of Lean practice in health care. It will look at the volume, type and origin of published academic work in Lean health care, describe the most common aims of Lean in health care and discuss the common "failures" of Lean interventions, and the causes for this. Following on from this – suggested success factors will be discussed for more successful interventions.

Lean health care remains a niche domain of academic publishing. In 2014, Lawal called the amount of research "limited" [140], though The rate of publication, has risen substantially since its genesis in the late 1990s [141] with the first paper being identified [142] as the work of Preston in 2000 [143]. In a systematic review, Anthony et al. [142] found that since the year 2000 there have been only 101 papers published in the field. Mason et al. in an earlier publication, found a similar number of papers published in the years during which their reviews overlapped [144] whilst Radnor & Burgess found a slightly higher number of papers, but not disproportionately so [145]. In the first 5 years of the millennium, only three papers were found, however, the rate of publication has accelerated rapidly towards 2016 (the last date of inclusion for material in Anthony's team's study).

Globally, the bulk of publications in Lean health care originate from three countries, namely the USA, the UK and Sweden [142]. Between them, these countries account for 62 % of published work that was found [142]. In the USA, Lean health care work is broadly centred around the Joint Commission Institute with the list of authors, organisations, institutes and facilities being very diffuse. In the United Kingdom, there is a notable axis of work taking place in the NHS Institute for Innovation and Improvement [142] and a particular focus by a number of academics mainly associated with the universities of Cardiff and Warwick. In Sweden, up to nine out of ten hospitals are claimed to have "gone Lean" to some extent [146], however the academic work is largely centred around the *Karolinska Institutet* the institute that among others, hosts the Nobel committee, that awards the Nobel prize in Physiology and Medicine [147].

Lean interventions have been seen in most areas of the health system. From pharmacies and supply-chains [148-150] to emergency departments [151, 152] to surgery [144, 153, 154] to documents management [155] and numerous others. Most commonly, Lean interventions attempt to reduce patient waiting times and lengths of stay [151, 156] or Improved infection rates [153, 157, 158] and Operating Room efficiency [153, 154, 159]. Adopting a systems thinking lens, it is important that these metrics be related to system design, and not "solved" without systemic remedies.

2 - 6.1. Quality of research

Moraros et al. call the body of work in Lean health care into question. In a systematic review, they found that Lean health care reported no statistically significant positive impact on satisfaction or outcomes, further they found that interventions tended to be expensive and were a burden on staff happiness though finally, they reported that some studies showed positive outcomes in process and flow [30]. This led to their concluding finding that "while some may strongly believe that Lean interventions lead to quality improvements in health care, the evidence to date simply does not support this" [30 Pg. 1]. This is a strong finding which they use as a basis to recommend that the quality and rigour of research must be strengthened to raise credibility of the field [30].

Although Maike et al. highlight examples of unsuccessful lean initiatives [160] generally there appears to be a publication bias, with unsuccessful interventions being rarely published [161, 162]. This notwithstanding, Andersen et al. lament that even where reporting is positive, the nature of the studies, and the shallowness of supporting evidence [163] makes claims of success questionable.

Study designs contributed to the uncertain outcomes. Broadly many studies were structured as single case studies [142] with greater emphasis on describing (often anecdotally [163]) the intervention than measuring outcomes (with an emerging qualitative bias [142]).

Moraros et al. show that only 22 out of 48 papers passed a methodological test [30], no papers used "high quality" methods and no use was made of control samples [30]. Follow up of findings more than two years after the study was only reported in two papers – calling into question claims about sustainability [30].

Glasgow's study [164] supports this concern finding that in 15 out of 47 cases, no meaningful data were used to assess baseline information nor to contrast the post-implementation success – and only three publications revealed data from more than two years post-implementation. This supports Anthony et al.'s claim that evidence is generally "sub-optimal" [142]

Largely, studies feature data collection at a distance, with approximately half of published work making use of interviews and questionnaires [142]. This is not in itself problematic however it could limit "buy-in". Relatively few studies report on direct observations and similarly data analytics are under-reported. Whilst methods like Meredith et al.'s analysis of operating theatres [159] could provide stronger empirical evidence for data driven decision making, such analytical papers are in the minority.

Given methodological and data gaps, Brandão De Souza claims that the interpretation of project success is generally unclear [165]. Alexander & Hearld find outcomes generally difficult to attribute to interventions, and therefore generalisation from work is difficult [166].

These studies highlighted that the constraints and limitations of the context need to be better understood to ensure that Lean achieves its promise [141] and does function as a cure (or a panacea) [167] for the woes of the health system, rather than the unsuccessful paradox that it presents [167] being implemented too narrowly, with lacking depth of understanding and nuance.

2 - 6.2. Aims

Lean in health care aims to strengthen systems and process capability [161, 168]. Lean interventions generally attempt to "bridge the quality chasm" [163 Pg. 2], or setting the objective of "moving from the health care that we have to the health care that we should have" [163 Pg. 2]. Radnor proposes that interventions should proceed over the paths of sustainability of change, organisational change and organisational learning [167].

Lean can be used to do so, by strengthening process capabilities [161, 169]. This is done, because Lean is seen as a valid means of addressing the complex and underperforming nature of health systems in a way similar to the gains that have been achieved in the automotive sector [116, 141, 142, 163, 167, 170].

A few key themes emerged from a spectrum of systematic reviews [140, 142, 144, 153, 161, 163, 171]. The aims for Lean implementations in health care were extracted from these sources. To structure these aims for this thesis, Radnor's model of "People", "Process" and "Sustainability" [167] is completed by the addition of a "Customer" dimension.

- People
    - Transforming organisational culture.
- Process
    - Improving processes and gaining operational efficiency,
    - Standardising and streamlining processes.

- **Sustainability**
  - Outperforming the competition and gaining a commercial advantage,
  - Eliminating waste,
  - Eliminating non-value-adding tasks,
  - Reducing staff and administrative inefficiencies.
- **Customer**
  - Providing better services,
  - Improved patient satisfaction,
  - Reducing delays and operational time,
  - Improving service quality.

Unfortunately, it appears that the promise of Lean health care is not always realised in its execution [141, 167, 172]. This means, that despite best intentions, Lean implementation is often lacking, unsuccessful or indeed possesses the data poverty that makes conclusions about the quality of the intervention questionable [163, 165].

2 - 6.3. Success of implementation

Fujimoto proposes a three-level hierarchy for Lean success, namely "process", "learning" and "evolving" [173]. These steps refer to the incremental maturity of how Lean is used, starting with the first level: "process interventions", simple, non-analytical use of tools. These are perhaps the "tool heads" of whom Seddon is dismissive [44]. The "learning step" is the layer in which organisations start to make use of standard analytical tools, implement change and have a means to "remember" solutions. The "evolving step" speaks more to Hines et al's [174] strategic or philosophical take on Lean, developing, innovating and growing on the basis of a strong foundation of an understanding of the philosophical underpinnings of Lean – targeted towards system-wide embedding of best- and innovative-practices.

Fujimoto's framework can also be seen through the lens of Jones' wording – "improve", "redesign" and "rethink" [170]. In that framing, Mazzocato et al. find that there is no evidence of mature operations on the evolutionary (rethink) level [175]. This shows that the implementation of Lean in health care remains largely at an operational level, tool driven and does not have the strategic energy that is required as per Hines et al., to create long-term, systemic and sustainable improvement.

2 - 6.4. Constraints

To further understand why implementation is perhaps less successful than hoped for, based on the seemingly self-evident usefulness of Lean in health care, it is worthwhile to explore the constraints that govern the system. In this regard, study design, level of execution and the nature of health care as heterogeneous [174] and complex [176, 177] will be explored.

Lean interventions are rarely executed at the enterprise level [145, 165, 178]. This means that implementations are bitty and, with little exception, not institutionalised [118] this differentiates it from Toyota who have ensured that Lean exists in the totality of their value chain [15, 170]. This leads to limited technical implementations that do not reach across the organisations [163]. This contrasts with Jones' encouragement that hospitals should not be seen as departments, but rather as a collection of value streams [170]. Though there is little evidence of any consolidation on this level (with an exception being the post-merger University Hospital of North-Norway, that did make use of some Lean principles, and considered value, merged departments and emphasised flow [179]). Without a system-perspective, lacking broad restructuring, Lean interventions may remain bitty [180] which is potentially detrimental to the greater health system [145].

Despite claims about low Lean maturity, a positive sign towards enterprise level interventions was reported by Burgess & Radnor [178]. They presented the results [145] of a large longitudinal study in the English NHS, making use of content analysis from publicly available sources [178]. From this study, the encouraging result emerged that the implementation of Lean in health trusts was shifting away from tentative interventions or islands of intervention towards systematic programs attempting to alter culture fundamentally.

A further constraint to the quick implementation of Lean ideas to hospitals is the complex nature of health systems. This notion of complexity is a broad field that includes the fact that as a socio-technical system [181] health care brings together the human elements, societal drivers and technical questions that range from machinery, technology and infrastructure, through to legal and financial frameworks within which the health system exists. This socio technical complexity which falls into Walshe's category of Heterogeneity [182] makes systemic leanness a challenge, and several authors highlight the complexity of health care, as possibly the greatest impediment to good Lean [142, 160, 172, 183].

In addition to these constraints Anthony et al's work has identified twelve broad categories that limit Lean implementation success. These are also read together with the work of Radnor & Walley [115] and have been summarised for this thesis into the following categories: resistance, incredulity, instrumentalism and framing.

2 - 6.4.1. **Resistance**

Anthony et al. found resistance from management and from staff. Lean remains an unfamiliar approach in health care and being required to change is an uncomfortable proposition for most people [142]. This is exacerbated by the generally slow pace of projects and the long lead time to positive outcomes [142], leading to participant demotivation[142], which may be worsened further by the sense that people are underpaid and overworked [115].

2 - 6.4.2. **Incredulity**

Andersen et al. paint a picture of the clinical mindset in health care [163]. It is strongly driven by the need for evidence and empirical proof prior to acceptance. The Lean approach is thought to lack sufficient, believable empirical evidence for its claims of providing gains, and therefore lacks the ability to persuade clinicians, who tend to be resistant to change based on what is perceived to be weak evidence [142]. Anthony et al. claim that there is a dearth of persuasive theory to back up an implementation and so resistance to participate grows. Radnor & Walley further identify a lack of real understanding of variation, systems thinking and the principle of flow as a basis to contextualise Lean [115].

2 - 6.4.3. **Instrumentalism**

The use of tools – instruments, is something against which Seddon cautions in his inimitable way: "beware of the toolheads" [44] he cautions; tools should not be used for the sake of using them [115], yet – this remains one of the major concerns for implementation. Anthony's team found that a lack of knowledge of tools and techniques, together with a lack of understanding which tools to use, is frequently cited as a reason for the failure of Lean interventions [142].

2 - 6.4.4. **Framing**

The management of Lean projects is generally poor. This leads on from Spear's concern that Lean does not get institutionalised as it has in the automotive sector [118]. This means, that there is weak coordination within a small project and across several small projects in a larger system. Equally, without clear framing, there is no agreement on the criteria by which the value of a project will be measured [142], nor is there agreement on an approach that should be used to address systemic complexity [142, 172]. Without a clear customer focus [115], the purpose of interventions is often concealed in seemingly uneconomic returns. This is further illustrated in the harm done though silos and too many targets[115]. This means that interventions are not focused and an overall lack of direction. Without proper framing, projects are likely to be less successful than they should be.

2 - 6.5. Factors for success

Although successful Lean projects in health care may be less common than expected, there have been numerous great examples where successes have been recorded. Andersen et al. [163] have identified the key drivers (in their words “facilitators”) for Lean success. Broadly, these speak to the constraints mentioned in the preceding section [142]. The key facilitators for success are management, and their involvement, the existence of a supportive culture, training, the availability of good reliable data and including clinicians in all initiatives.

Management’s role includes the need for commitment to the project and a coordination role to ensure that the implementation has systemic strategy and support. The second facilitator speaks to a supportive culture, that encourages experimentation and participation in improvement projects, does not exclude participation and acknowledges gains. This should be backed up by the third facilitator: training. Training should equip teams with the necessary knowledge, understanding and insight, to bypass failed instrumentalism and to equip all with necessary insights to overcome incredulity. Data quality is a crucial facilitator of success, and Andersen et al. emphasise that successful implementations derive from successful data strategies which in turn also leads to evidence based confidence in Lean [163]. The final important facilitator for Lean success is the inclusion of physicians in the projects.

2 - 7. Capacity and Demand

The Vanguard method advocates that any improvement requires that the demand (as viewed from the perspective of the customer) in the system is understood first [32, 44, 184]. This view is shared by Marr & Neely [185]. Walley shows that the (mostly private) service-literature encourages demand driven approaches, meaning that systems should be designed after answering the question: "what demand do I have?" [186 Pg. 880]. Leading on from this, the system can be appropriately designed, with an ability to respond to the demand placed on it [186].

This highlights the importance of understanding how demand comes to be, what its constituents are and how it is measured. Once the demand is understood, it is possible to design the system and the work to achieve the desired outcomes [32].

2 - 7.1. What is capacity?

Capacity is used as a term by economists to describe the ability of a supplier to provide goods or services to customers [187]. Capacity can exceed the demand placed on the system by customers, which Mankiw calls "Excess capacity" [188]. This excess capacity is generally not made use of, as prudent organisations will not produce goods or services in excess of what can be sold [188]. Nevertheless, this capacity exists hypothetically and may be realised when demand rises for whatever reason.

In services broadly, understanding capacity is perhaps more challenging than in pure manufacturing environments. Where goods are produced, a quantum model of capacity can be applied – for example, based on machinery, production rates and overall efficiency, a firm may have a capacity of producing 100 000 ball-point pens per day.

This is not so simple in services where the nature of the product varies from case to case [137] and as a result expressing capacity in terms of 'units of production' [44] becomes less exact. Whereas manufacturing can express their capacity in terms of completed goods – services, and especially a complex service, like health care [176, 177] – must find proxies to represent their capacity.

Several proxies are used to describe health care capacity. Demir et al. express the capacity of a hospital in terms of physical infrastructure (mainly beds but also operating rooms and outpatient clinics) and human resources (mainly clinal personnel) [189]. Allder et al. critique the idea of beds being the apex descriptor for health system capacity. They point out that due to demand variation and certain standard NHS practices, including infrequent ward rounds, in many instances lack of beds capacity is "self-inflicted" [190 Pg. 15]. Instead they favour a systems-view that allows the capacity to rise by managing and balancing demand rather than pure capitalisation models [190].

As previously mentioned, because services broadly, and health care in particular, cannot prepare and buffer services during quiet periods [112, 191] services can only be provided when required, meaning that capacity has an instantaneous element to it. This means that quiet times do not necessarily imply excess capacity, but equally, when the system is over-run (Mura and Muri) [63, 141], this does not necessarily suggest that the system-capacity is inadequate [192]. To make such a conclusion, a long-term view must be taken.

Given that capacity in health care is firstly difficult to define, and probably inaccurate in complex projects in capacity management, a well-reasoned definition is necessary, but this must be accompanied by a willingness to modify the definition if the context requires this. This thesis is not a thesis in the domain of capacity management, though from time to time the system capacity will be discussed. In such cases this thesis uses the definition for capacity proposed by Walley et al. [193].

Capacity is "value added activity in a fixed period of time that a service can consistently achieve" [193 Pg. 11].

Unpacking this definition, important elements emerge. These include that non-value-adding activity does not classify as a component of the system capacity, which is consistent with Hopp & Spearman, in their thinking about detractors [40]. Similarly, capacity must be time based, thus emphasising that pure hypothetical capacity must be avoided in favour of the actual efficient ability of the system to provide the service. Finally, the notion of consistency is key. Outlier peak capacity must not be the system's ability to provide the service, but rather the "centre-line", repeatable, capacity must be considered to set a real expectation of the level of service that can be provided.

2 - 7.2. What is demand?

In classical economic terms, demand is the amount of a product, or a service, that customers are willing to purchase at a certain price [187]. Unpacking this definition reveals the importance of the choice of the customer. Implicitly the customer sets the value of a good through their willingness to consume it at a given price [188]. This further shows that demand is a phenomenon that is generated through the choice or preference of customers. Demand does not exist in isolation of a need in society that seeks to be fulfilled.

Demand is what the customer wants or needs. This simple description adequately describes how demand works, however, to move from the notion of demand towards practical means to measure and record this phenomenon is a more complicated problem.

It follows that demand places a load on the system. Demand translates into what the system needs to give to satisfy the customer's need. This need is varied and not standardised in services as it is in the manufacturing industry [44]. The recipient plays an important role in whether a service has satisfied their need [121] – whether the service has achieved the nominal product value [173].

2 - 7.3. How is demand measured?

Knowing which units are used to describe demand is an important element of understanding a system [44]. In the academic literature, a variety of units have been used to describe demand. Whereas in the manufacturing industry, demand can be discretely measured in terms of production. As an example, car manufacturers need to make a certain number of vehicles (pulled from customer orders [14, 15]). Thinking in terms of "units of production" however, translates poorly into the service environment [44, 121].

Production-unit thinking leads to viewing demand as the number of patients seen [13] or the number of calls taken [194] or the number of people in a queue [13, 19, 149, 152, 195-197]. This is a simple measure that requires nothing more than counting the individuals in a queue or who have received a service. The simplicity of doing so however neglects to consider whether each of these demand instances represents equivalent demands on system resources. In almost all cases, one individual in a queue requires a unique amount of work from the system and is therefore not equal to any other demand instance. Or, as Seddon expresses it, because the product in services varies, from case to case, it is difficult, or impossible, to standardise product or process [32, 44, 137]. Therefore, measuring service demand by "production" units is not accurate and often misleading [32, 44, 137].

A clearer sense of the demand from an individual, is the amount of work that the system needs to do, to achieve the desired outcome that is described by the unique product sought in that service interaction [44, 121]. It is not about how many people a service system sees, but rather about how much resources are required to meet this demand [98]. Or how much capacity is depleted in servicing the customer's need.

2 - 7.3.1. **Lean consumption**

This thesis takes the position that demand is a consumption model [121, 198], a view favoured by Womack & Jones in service industries [121, 122]. This thesis takes the view that demand is the vehicle by which the resources of the health system are depleted. The resources that are consumed, or anticipated to be consumed (i.e. queues) include, but are not limited to, the skill and time of medical professionals, the physical infrastructure, and other resources such as consumables, equipment, money, information, laboratories and medical records [199]. The list of resources in question differs from case to case. With a systems-view, these resources are unlikely to be consumed in isolation of some (or all) the others. The demand, therefore, is equal to the sum of resources required to achieve the desired outcome.

A weakness of the consumption model is that it considers *consumed* resources. This implies that cases where a need is never satisfied, for example a patient who dies before receiving bypass surgery, does not count towards demand. For that reason, this thesis takes the view, that demand is the real consumption that the system experiences but also hypothetical consumption of health system resources, which would include queues (consumption in waiting) and unsatisfied demand (consumption denied). This is an important consideration, because the first step in improvement is to understand demand [44, 184], however if the work that the system didn't perform simply "disappears", demand is not understood, and an adequate response cannot be engineered. This means that strategic system designs as favoured by Hines [174] will inadequately respond to the needs of society.

To define demand so precisely; in terms of every fraction of utility extracted from each resource leads to an ungainly, and practically unimplementable measure. For that reason, this thesis takes the broad view, that demand can usually be expressed in terms of the time consumed, as this is broadly a decent proxy for the amount of work required and by extension, the resources required to complete a service.

Given that service systems are generally delivered in person [112] the key resource consumed during the provision of a service, is the time of the person delivering the service. Of course, health systems are too complex to reduce them only to individuals, and at times, the capacity of the system, and therefore the ability of the demand to consume resources, is defined by a different resource, say MRI machines. In general, most demand reduces to the people delivering the service, or the infrastructure required to perform the service – the 'doctors and beds' model [189]. Nevertheless, in all cases, discretion is required to realise that the key unit of consumption must be derived from the context.

Although the consumption model is largely used in this thesis, in Chapter 4 this thesis proposes a model for more complex demand measurement, that considers the consumption of time as well as the consumption of complex resources, with a proposal for how this complexity may be assessed. This differentiates between the level of speciality required, or between an hour of porter's time and an hour of super-specialist's time (and a continuum in between). This model may be useful for strategic system assessment leading to the ability to design system capacity based on an understanding of demand [44, 174].

2 - 7.3.2. **The gap between capacity and demand**

Poor service delivery can often be traced back to an inability of the system, to meet the needs of those who make use of it. Typically, in an environment where demand outstrips capacity the good will trade at a premium and speculators will enter the market. By contrast, where capacity outstrips demand, the market becomes highly cost competitive and the participants that cannot compete (often on price), will typically exit the market. [70, 200, 201]

Where capacity shortages exist, a gap emerges between capacity and demand. Such a gap results in the inability to serve all patients in a timeous manner with satisfactory quality. This means that patients will necessarily accumulate in queues, which will become longer. The higher demand places an increased burden on the service provider who is probably more prone to make mistakes leading to poorer quality of service in general. If additionally, market forces are permitted to play a role in this relationship, the price would also rise according to the laws of supply and demand [201], however in health care this is not always the case, because public health care prices are often legislated, fixed or regulated.

2 - 8. Closing the gap between capacity and demand

This section explores how the gap between capacity and demand broadly, and in health care, more particularly can be addressed. Although the classical investment approaches are mentioned, industrial engineers tend to focus on closing the gap through alternative means – primarily through improving system efficiency [65]. Sasser speaks of two approaches to closing the gap, the strategy of "chasing demand" [112] which aims to find capacity to meet demand, or levelling capacity, which is a strategy to managing demand, either through reducing it, or by moving it into a more balanced shape [112].

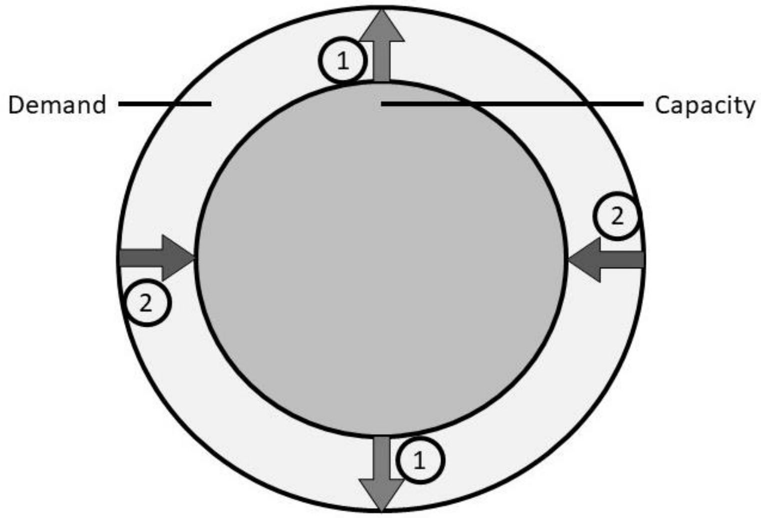


Figure 2-6: Conceptual schematic of capacity – demand gap.

The gap that emerges when demand exceeds capacity is shown conceptually in Figure 2-6. Notice that the circle representing demand is larger than that of the available capacity. The result of this is a gap and closing such a gap can be achieved in several ways. There are two macrooptions for addressing the gap. Firstly, capacity can be increased, which can be systemic (provide or control strategies [199]) or transient (chase demand [112] or match [199] strategies). This is symbolically shown by ①, or demand can be managed (influenced [199] or reduced (as shown by ②). In both cases, the gap will narrow.

The easiest approach to addressing the gap between capacity and demand is by increasing capacity. Capacity can often be increased through capital spend [112] – which may be expensive, but does not require specific skill, through additional resources, scrutiny or efficiency to be achieved. It is noted that in resource constrained environments, spending money may not seem easy – however this remains easier than restructuring the system to increase capacity.

2 - 8.1. Increasing capacity

It is proposed that capacity can be described by the following relationship:

$$C_a = C_t \times E \tag{4}$$

With, C_a is the apparent capacity, i.e., the current capacity of a system to serve patients, C_t is the true capacity, i.e. the design capacity, of the system. E is the efficiency of the system.

The traditional approach to closing the gap between capacity and demand, is by raising capacity to meet demand more closely. capacity can be raised in two ways; firstly, by expanding the system, which generally requires capital (increasing C_t in Equation 4) or secondly by improving system efficiency which often can be done at little or no cost (by raising E in Equation 4) [13, 197, 202, 203].

2 - 8.1.1. **Raising actual capacity**

Increasing actual capacity is most easily achieved by investing in new infrastructure, building more facilities, raising budgets, and employing more people. This approach is followed in the strategic documents published by the South African department of health, which prioritise spending as the remedy for closing the gap and improving service delivery [204, 205].

Closing the gap by capital injection however neglects the idea that the apparent capacity is not in fact the true capacity of the system, but merely the capacity that the system is capable of because of inefficiency. This means, that expanding an inefficient system may raise capacity but also expands inefficiencies.

To improve capacity, the lowest cost approach for improving output and productivity (service delivery) is to explore areas in which waste can be engineered out of the processes [65, 90]. In general, such interventions can be introduced at little, or even no, cost with dramatic results [13, 152, 197, 206]. However, as Section 2 - 6.3 showed, poorly run Lean initiatives are often unsuccessful.

2 - 8.1.2. **Improving efficiency**

The apparent capacity is positively associated with efficiency, this is shown in Equation 4. This is important as it serves as the underlying rationale for improvement projects in health care (and elsewhere). Recalling Figure 2-6, raising efficiency has the effect of inflating the dark grey capacity-balloon, thus narrowing the gap.

The author's previous research interest lay in raising efficiency in health systems, and the following two examples show how this was done.

Operating Theatre Utilisation [207]

In a major hospital's operating theatre, simple modification of the list-scheduling system resulted in the utilisation levels increasing from roughly 45 % to 85 %. This was achieved by modifying the list scheduling philosophy and shortening the individual surgeon time allocations. It was found that most lists were incompletely consumed. In fact, of the standard six-hour allocation, only 14 % of lists exceeded two hours, and 98 % of all lists were completed within 4! This wastefulness could be eliminated by a cost-free management change resulting in major utilisation improvements [154]. Secondly it was discovered that poor discipline led to delayed starting times that cost an average of ninety minutes for every operating theatre every day. The cumulative cost of late starts was dramatic, and the time wasted considerable. Alerting the system to ensuring that punctual starts are managed, allowed for more effective use of the spaces.

Medical Trauma Department Efficiency [152]

In other work, the author explored the efficiency of a medical emergency department. Here it was found that the queueing time could be materially improved by simply addressing two key issues, namely shift structures and facility layout.

The shifts were structured in such a way that the greatest number of medics was allocated for the times of greatest queue lengths. The error in this, is that the queues accumulated because doctors were not on duty at the time of greatest arrival rate. And the high queue length was a consequence of the accumulation of patients during this time. Restructuring the shifts to correspond to the time of highest arrival rates led to reduced accumulation and ultimately shorter queues but also limited Muri and Mura (overburden and unevenness on the medical professionals.

Before the study, the layout required 700 m of movement for patients in seeking treatment. These journeys were rarely done alone, and a physician nurse or porter usually moved with the patient. The overall loss to the system due to this was considerable and a revised layout reduced travel distance to 250 m which in turn meaningfully reduced overall cycle time. Simulated reductions in queueing time showed an improvement from six hours and fifty minutes to three and a half hours. [152].

These case studies highlight the gains that can be made by employing a Lean approach of health systems and making process and system improvements. Although efficiency interventions are useful and common for narrowing the gap that is shown in Figure 2-6, the burden placed on the system through unmanaged demand is thought to be significant.

An alternative approach to raising capacity to meet demand, is to reduce demand to converge to the available capacity [186]. To do so however, requires an understanding of demand; how it is constructed and what the underlying mechanisms and drivers are [44, 191].

2 - 8.2. Strategies to manage demand.

Services ought to be demand driven, this is often not the case, where they are frequently resource- essentially capacity- driven [186]. The idea of being demand driven, links neatly with the Lean pillar of "pull" [14, 65] – or the idea that the system should produce (the service) at the rate of consumption of the customer [65].

Sasser introduced the "chase demand" and "level capacity" strategies to deal with the gap between capacity and demand in service systems [112]. Walley & Adams contributed [Redrawn from: 193 Pg. 14]. This image captures the essence of the two philosophies. Chase demand models are responsive to demand dynamics and adapt to system load (Crandall & Markland refer to this as the "match" approach [199]). The level capacity strategy keeps capacity steady and allows demand to be dynamic around this level. Where possible, peaks that exceed capacity are addressed by a variety of techniques, if possible.

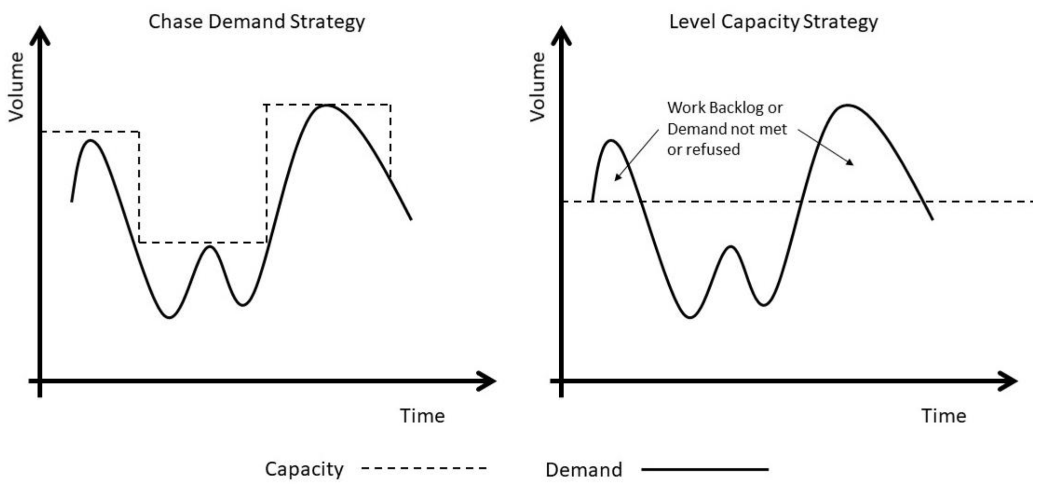


Figure 2-7: Chase demand vs level capacity strategies [Redrawn from: 193 Pg. 14]. [193 Pg. 14]

The chase demand strategy is designed around the belief that capacity can be rapidly added to respond to (largely transient) demand changes. Such rapid responses to the market generally require that the type of work that is to be performed possesses low complexity and onboarding new staff is simple and quick as in the case of a hamburger stall when a large football match takes place.

Unfortunately much of the service industry (and health care especially so) is, as Gupta et al. emphasise, "a knowledge work intensive industry where deep technical knowledge is the baseline skill [111 Pg. 1045] which violates the conditions [112] for being able to "chase" effectively. When chasing, the system must rapidly adapt to demand fluctuations, by scaling the capacity in response. To achieve this however, the skills requirement should be low as the capacity that can be added swiftly is generally of low technical skill which is poorly paid, has a low level of training and Sasser himself refers to the working environment as a "sweatshop" [112 Pg. 135].

Crandall & Markland propose four strategies for managing capacity and demand [199] which are refinements of Sasser's models [112]. Their strategies are match, provide, control and influence. Match-strategies design capacity to be equal to demand, which tends to be a high-cost option, however, clearly provides the highest level of service to customers. The influence strategy seeks to move demand to create a levelled demand, whilst the control and provide strategies seek to provide a fixed capacity, either at the average demand or the peak demand, respectively.

The level strategy by contrast does not have systemic response systems to deal with the occasions when capacity is inadequate. Therefore the system must either be designed to respond to average load (which Crandall & Markland call "the control" strategy, or the system can be designed for demand peaks ("provide model" [199]). In the instances where peaking work can be shifted to times of lower demand, this is an effective demand management strategy that can work in health care [112] which Crandall & Markland refer to as the "influence" strategy [199].

A common interpretation of the peaks exceeding available capacity, is that the system is inadequately designed to respond to demand, this needs a broader view to assess. Allder et al. propose that capacity is only inadequate if the queues waiting for service continuously grow [Redrawn from : 192 Pg. 32] as can be seen in Figure 2-8.

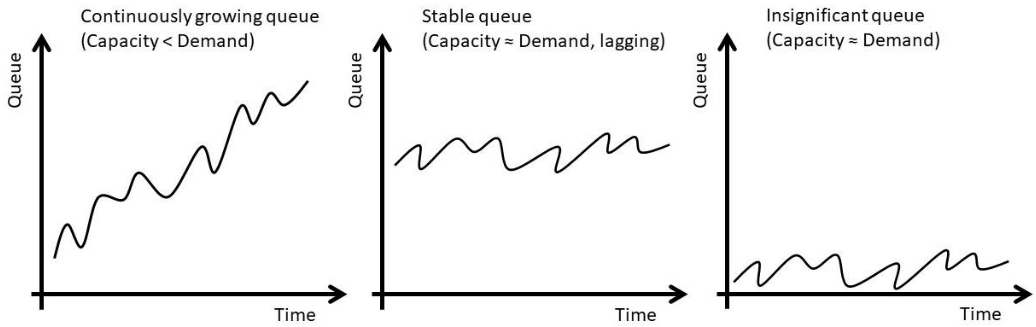


Figure 2-8: The relationships between capacity and demand and queues [Redrawn from : 192 Pg. 32]

In the instance where waiting times exist, but the queue is not growing, then demand and capacity are (approximately) matched, but with a time shift. This capacity-lag- [192] (influence- [199]) strategy is a means by which the periodic excess capacity shown in Figure 2-7 can be used to compensate for the peaks. This lag strategy is a viable system design because systems can be capacitated at lower cost, than by incurring the complexity and cost to match capacity and demand.

The strategies presented here assume that the demand that is being experienced is the demand that the system should experience. This means that in the absence of questioning the fundamental construction of demand, the demand must be serviced in full, or systemic compromises must be made to decide which demand not to service.

This thesis echoes the belief that understanding the constituents of demand allows for the discovery of demand that is not demand, but rather wasteful, and therefore represents an opportunity for systems improvement [194].

2 - 8.3. Practical approaches to managing demand.

The broad strategies for demand management include shifting demand around [208] which Klassen and Rohleder expand on as changing the timing of demand or re-channelling the demand to other resources [209, 210]. Heineke adds the option of curtailing demand as an acceptable strategy for managing demand, which is important, because demand in health care is often treated as fixed [190]. Whether demand truly is fixed should be explored in greater depth as actions could be taken to reduce demand. Key demand management strategies include clinical and system changes.

2 - 8.3.1. Clinical approaches

The idea of demand management in health care was raised in 1995 by Vickery & Lynch [19]. Demand can be significantly reduced by reducing disease [211]. This can be achieved in several ways, including a stronger primary health care system [204, 205], improved health information services and community health work [212] and other preventative and prophylactic interventions [213].

2 - 8.3.2. **System approaches**

Demand can be discouraged through pricing. This means that where cost attaches to the patient, demand decreases. For example, in Bangladesh, Out of Pocket expenditure for health costs is approximately 64 % [214]. This leads to economic exclusion of poorer patients [215] and the net effect is that the load on the system is considerably lower than the actual disease burden in society. As a result, Bangladesh has the highest level of catastrophic health care expenditure in the South Asia region [214], and endangers its ambitious aims of universal health coverage through economic exclusion [216]. This economic exclusion is not a phenomenon only of the least developed countries [217] but as is seen throughout this literature review, also affects the most prosperous nations on earth [8, 9, 218].

Conversely, where patients are shielded from actual health care costs by third party insurance companies, demand increases uncontrollably [219]. In the USA, Health Maintenance Organisations (HMOs) play an important role in taking pressure off hospitals. An HMO is a health insurance that allows access to networked primary health services, on demand, for no, or low, patient payments, variable based on patient risk profile [220]. Several studies identify the role of HMOs as reducing demand. However, this interpretation neglects to realise that this demand is only moved, and still acts on the health care system, albeit at a different point [220-223].

A stronger public health care system shifts the disease burden downwards to lower cost and lower energy interventions [209]. Numerous other interventions should be strategically prioritised to take pressure off the health care system. The drawback of prevention over cure strategies is that the benefit is invisible[7].

A note on managing Covid-19's burden on health systems.

A topical example at the time of writing this thesis, is the case of Covid-19 – the disease caused by the SARS-CoV-2 virus [224]. This case showcases both clinical as well as system considerations. As this is an emergent issue, many of the reports and resources are sourced form grey literature such as reports by the World Health Organization (WHO) or The American Centers for Disease Control and Prevention (CDC), or newspaper and other press reports.

Covid-19 threatens health systems by overwhelming their capacity. In China – the original source of the disease [3], thirteen new hospitals were built within weeks and a further nineteen makeshift hospitals were planned to deal with the outbreak [225]. This demonstrates the asymmetric load that such an epidemic can cause. As the epidemic spread globally, it was declared a pandemic [226] and began to place a similar burden on the health systems of other countries. As the pandemic grew, the point was reached after which containment could no longer limit the growth of the disease. Germany's government reported that 60 – 70 % of the German population was likely to get the disease [227], whilst MIT reported that up to 214 million Americans could get infected, leading to as many as 1.7 million deaths [228]. The Centers for Disease Control predicted that up to

[7] Someone not getting measles is not an event, whilst being healed of measles is – so securing funding for prophylactic or other preventative treatments including primary interventions, may be more complex.

21 Million Americans might need to be hospitalised, which would be unmanageable.

The mass of cases makes preventing the spread of the disease no longer feasible, rather the focus shifts to preventing that everybody gets the disease at the same time [229]. To this end, global health agencies, such as the World Health Organization recommend several approaches to manage the spread of the disease. These approaches include managing personal hygiene – by washing hands and avoiding touching one's face, social distancing – in other words limiting interpersonal contact and infected people taking additional precautions to not spread the disease, such as the wearing of masks, isolation, quarantine and avoiding all physical contact[230].

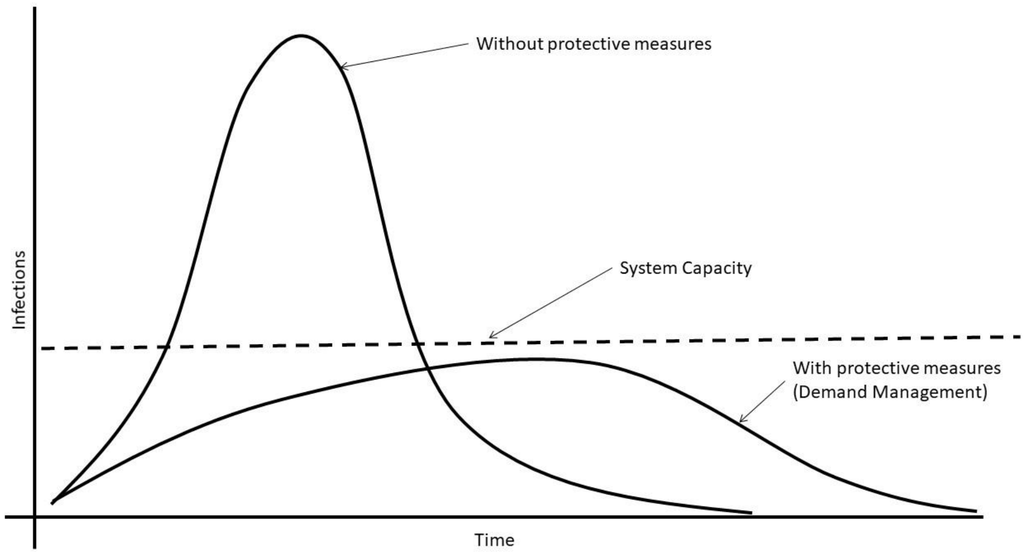


Figure 2-9: The effect of slowing transmission on system demand and response capability [229]

Figure 2-9 shows the effect of following good mitigative behaviours[8]. Slowing the spread of the disease is unlikely to limit the total number of people who do get infected and place a burden on the health system (the integral of the curves), but can manage the pace at which this occurs [229]. This leads to less over-burden (Muri) and unevenness (Mura) [141]. In other words, managing demand in this case does not mean reducing demand, but rather only focuses on *spreading* (or "influencing" [199]. demand more evenly to protect the health care system [229]. Of course, the idea of the system's capacity not only implies the ability of the health system to deal with COVID-19 cases, but all other pressures the health system would ordinarily face too [229].

The COVID-19 pandemic aligns to the purpose of this research, the realisation that health systems need to be treated as whole entities, that load affects the system and not just points. The pandemic further highlights the usefulness of addressing demand, and that levelling demand leads to improved overall health outcomes.

2 - 8.4. Ideas of how demand is shaped.

2 - 8.4.1. **Patient motivation shaping demand.**

Vickery & Lynch set out to understand what the constituents of demand in health systems are. They explored how these are formed, what the drivers for demand are, and finally what categories exist for demand [19].

[8] This curve has become very familiar during 2020 and is generally credited to the CDC, however the originating author was Qualls [229]

Four major demand modalities were identified [19], namely morbidity (unwell patients), perceived need (the unqualified decision by patients to seek medical care), patient preference (the expression of choice as to where when or how medical care was sought) and non-health motives (the consumption of health system resources for reasons unrelated to care). Their model can be formulated as:

$$D_t = D_m + D_d + D_p + D_n \tag{5}$$

With: D_t : Total demand; D_m : Morbidity demand; D_d : Perceived need; D_p : Patient preference and D_n : Non-health motives

Morbidity is an objectively simple category to identify; for example, a trauma case arriving at the hospital, an ill person, an injury or some form of infirmity. The other categories are more textured and therefore more difficult to identify. Perceived need is complex because medically untrained patients often seek care without realising that care is unnecessary. Patient preference has been researched by looking at decisions patients make through 'informed choice' or informed consent [19] which in some cases has led to reduced demand at a particular location [231]. This neglects a systems-view, that considers that an active choice loads one part of the system in preference over another. This means, that the overall system load remains unaltered. Non-health motives, are unrelated to the patient's state of health, and include purposes like sick leave administration, disability or worker's compensation benefits [19]. This category is difficult to measure and there is relatively little quantitative information on the impact of non-health motives on the load on health care systems [19].

To reduce demand, reducing the magnitude of any one of these elements would be useful. Emerging out of this, Vickery & Lynch proposed that the major areas that could be managed were the patient driven modalities, namely perceived need and patient preference [19]. This implicitly states that morbidity demand is fixed – which is a crucial departure point for this study, as this thesis argues that a portion of morbidity demand derives from failure demand, and that morbidity demand is therefore not fixed at the rate at which it manifests.

2 - 8.4.2. **System driven mechanisms shaping demand.**

As physicians, it is not surprising that Vickery & Lynch' perspective of demand patterns is based on patient behaviours and motivations and neglects systemic mechanisms that can create or shape demand [19].

With a Lean lens, service demand is often thought of in terms of Seddon's mental models [44]. His generic formulation for the service industry claims that demand consists of value demand (the demand we wish to encounter) and failure demand (that demand "which is caused by a failure to do something or do something right for the customer" [44, p.76], resulting in their eventual return).

Seddon's ideas can be formulated as:

$$D_t = D_v + D_f \tag{6}$$

With: Dt: demand; Dv: value demand *and Df:* failure demand

Whilst Seddon's classical view of demand may suffice for service organisations such as banks, local authorities, telecommunications, and other public services, it is necessary that it is adapted for the health care context. Due to the complex hierarchical nature of the system and the complex nature of disease [177], it is thought that health care is an environment in which Seddon's definition requires refining. Seddon's view holds that demand is either value demand or failure demand. However, there are examples from health care that fit in neither category comfortably. For example, a patient who places load on the health system yet is not ill. If no previous system failure caused their presence, they are not failure demand, but also clearly not 'the demand that the system wants to see' [44]. Therefore, some expansion of the categories is required.

Mental Health Scotland [232-234] have developed an expanded modularisation of demand that contains more demand categories to understand demand in their system (shown in Equation 7) . Their model consists of actual demand failure demand, created demand and hidden demand. The particularly valuable contribution made by this model is the idea of hidden demand, providing language for a potentially significant demand burden that is currently unobserved and as such represents a meaningful system risk but also introduces language to describe a portion of demand that must be considered in any strategic system designs.

$$D_t = D_A + D_F + D_C + D_H \tag{7}$$

With: D_t*: Total demand;* D_A*:Actual demand;* D_F*: Failure demand;* D_C*: Created demand and* D_H*:Hidden demand*

In chapter 4 this thesis develops a synthesised model that attempts to capture all the modes of demand, from the systems- as well as the clinical-view. At the risk of redundant repetition, that section is not reproduced here.

A feature in the models presented in this thesis that take a systems view is that of failure demand. Failure demand is a central concept in this thesis and although the descriptions so far have been at a higher level, the next section will explore this phenomenon in greater depth, describing failure demand, providing a synthesised visual model to understand the nature of failure demand. This will lead into a discussion on the impact of failure demand and will show previous work in failure demand in health systems. This will lead into a critical discussion of failure demand with a particular focus on Seddon's claim that "failure demand is entirely in the control of the organisation" [44 Pg. 76]. Which suggests the system-view he imagines for failure demand.

2 - 9. Failure demand

Originally, Seddon introduced the idea of "demand that we do not want" [136]. It took about a decade for him to find a term that satisfactorily described the concept. In his seminal work "Freedom from command and control"[9], Seddon introduces the term failure demand, and defines it as

---

"demand caused by a failure to do something or do something right for the customer" [44, p.76]. .

---

[9] Consider Fayol's idea that "Command, Coordination and Control" was the central role of leadership [75].

In practical terms, this means that failure demand manifests as the need to do work more than one time, often because of a mistake made in the system. Figure 2-10 shows such a system in which in addition to work being completed, a portion of work also exits the system incomplete.

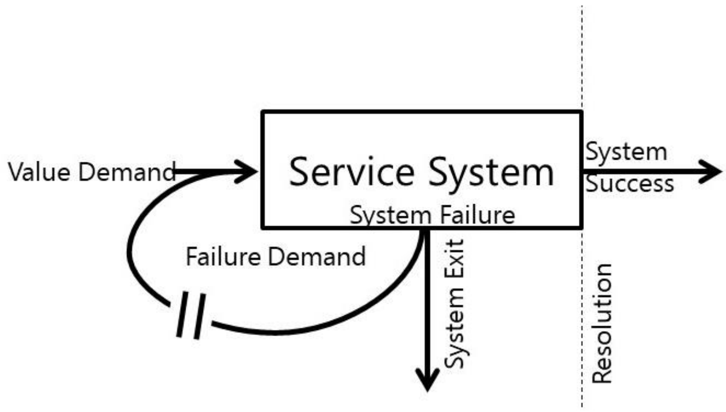


Figure 2-10: Generic system showing failure demand interpreted from [44]

The system receives a certain demand which consists of value demand and failure demand. This total demand is addressed by the system. A certain proportion of work passes through the system unimpeded; this thesis refers to this as system success. Some work elements however do not successfully transition through the system at various stages of the process and this failure occurs for a variety of reasons. In this thesis, this is referred to as system failure. The work element may return to the demand queue at a later stage after an unspecified time delay. The work elements that return, represent failure demand – the demand created by a previous system failure – or because of not doing something, or doing something wrong. This means that the total demand is increases by failure demand elements, which place the same (if not greater) burden on the system as they did previously.

2 - 9.1. Impact of failure demand

Seddon analysed call centres, where he found that failure demand accounted for as much as 80 % of demand and later used failure demand to highlight wastage in the public sector [44]. A further study [185 Pg. 14] observed that "it is critical to understand and classify the nature of demand … often calls are unwanted … by classifying demand … it becomes more manageable … unwanted calls could then be reduced or even eliminated". In a study looking at town planning document checking, "failure demand [was] a product of the system and it cannot be eradicated without changing the system" [235 Pg. 9], in this case, the failure demand contribution to total demand was roughly 40 %. This means that 40 % of documents appear on a planner's desk for a second time, having previously been sent out of the system.

The magnitude of failure demand in organisations varies widely. Piercy & Rich state that failure demand in a service industry is often between half and two-thirds of actual demand [110], whilst Seddon has found this value to be as high as 80 % [44]. Seddon goes on to claim that "failure demand is often the single greatest source of waste" [44 Pg. 39] in organisations, and that if it is predictable, in other words, patterns can be identified, the causes for it must be removed. This – he calls "the single greatest economic lever" [44 Pg. 277] available to managers to make improvements.

It is proposed that failure demand plays a significant role in health care demand, and it is surprising that little work has been published in trying to understand or describe the problem in the health care industry. A notable exception is the relatively recent work of Walley Et al. that looks at failure demand in UK primary care, and finds that staffing issues cause emergency service overload, chiefly near month end and at Christmas time [194].

A caveat about "Supplier induced demand".

Despite apparent semantic similarity, failure demand should not be confused with so called "supplier induced demand" which is a demand created by the health system which effectively takes advantage of patients' lack of knowledge [236, 237]. For example, if a doctor (whose motives may be questioned) recommends an unnecessary Caesarean delivery, that demand is in fact created by the supplier of the demand. Similar examples include clinically unnecessary imaging scan, or a clearly uncalled for blood test. This behaviour is called "supplier induced demand" and raises system demand - however, it is important not to mistake it as failure demand.

2 - 9.2. Some cautions about failure demand

This thesis set out with the intention to explore failure demand in health systems. It became evident however, that the definition of failure demand proposed by Seddon has problems and could be refined or redefined to fully address the difficulties primarily associated with the directionality of failure [44], the event horizon of the system [44] and the tight binary understanding of demand as proposed by Seddon [44]. For this thesis, three key words in the definition for failure demand are placed in question – namely "for the customer".

To bridge this gap, the contradictory binary demand landscape is augmented with other, observable modalities. The thesis describes a broader case of failure demand, that includes the service recipient's agency in producing failure demand by stretching the system boundary to include all actors.

2 - 9.2.1. **Binary demand**

Seddon's thinking on demand is powerful. It categorises demand into one of value demand, that demand that we want to experience, and failure demand which is the demand that we do not wish to experience [136]. All demand can be categorised into these two categories [44]. This binary classification is problematic, especially as it may be internally contradictory. Based on the binary nature of demand, it is reasonable to deduce that if demand is not failure demand, it must therefore be value demand (and vice versa). This is however not necessarily true, as healthy people do place a burden on the health system.

In this instance, it is a first encounter, the system has not done anything wrong, or not done something to cause this demand, but equally this instance is not value demand – this patient is not ill, seeing this patient, is not the demand that the system needs to see – and therefore this thought experiment shows a contradiction - describing a scenario in which demand does not fall into one of Seddon's binary classifications. To respond to this, it may be necessary to introduce other modalities of demand to augment Seddon's binary classification to be able to describe a greater variety of cases.

2 - 9.2.2. **Seddon's system**

To understand the way that Seddon views a "system" it is useful to reproduce a table that he used in "Watch out for the Tool heads" [119 Pg. 7] and in Freedom from Command and Control [44 Pg. 15].

Table 2-1: The difference between command and control and systems thinking [44 Pg. 16]

Command and Control Thinking		**Systems Thinking**
Top-down, hierarchy	**Perspective**	Outside-in, system
Functional	**Design**	Demand, value, and flow
Separated from work	**Design**	Integrated with work
Outputs, targets, standards, related to budgets	**Measurement**	Capability, variation: related to purpose
Contractual	**Attitude to customers**	What matters?
Contractual	**Attitude to suppliers**	Cooperative
Manage people and budgets	**Role of management**	Act on the system
Control	**Ethos**	Learning
Reactive, projects	**Change**	Adaptive, integral
Extrinsic	**Motivation**	Intrinsic

Table 2-1 shows how Seddon's view of systems thinking is that knowledge and learning should not be separated from work [44]. Yet, as he points out, modern management is separated from the work, "detached, remote" which is in his view, "the primary cause of poor economic performance" [44 Pg. 16].

That sets the fundamental principle of systems thinking for Seddon – the notion of outside-in; or in other terms, seeing the organisation (and products, and value, and performance, and quality) from the point of view of the customer [44]. This view does not substantially diverge from Ohno

[14] or Womack & Jones, neither for manufacturing [65] nor for Consumption [121, 122]. "Command and control" organisations by contrast design processes from the appearance of economies of scale and from a functional point of view [44]. This approach is often chosen for the sake of convenience of the organisation, but rarely leads to systemic benefit [137].

2 - 9.2.2.1. The system-boundary of failure demand

Recall that Seddon defines failure demand as: "demand caused by a failure to do something or do something right for the customer" [44, p.76].

This formulation reveals an important insight into the way that the system boundary is drawn for service systems. Looking at the definition above, the last three words: "for the customer", capture an important nuance of the way in which Seddon views the system that causes failure demand. "For the customer" implies a vector – a directionality in terms of the ability to cause failure demand. Seddon supports this interpretation immediately after this definition by stating that "failure demand is entirely in the control of the organisation" [44, p.76].

This suggests that the relationship between service provider and service recipient is one in which there is a directional ability to cause failure demand; the organisation can, the recipient cannot. Figure 2-11 shows an interpretation of the way in which the system is viewed based on this formulation. As can be seen – Seddon's system boundary (the course dashes) is drawn around the organisation, and that there is a directional relationship between the organisation and the service recipient. This causative relationship means that failure demand is created in this system from the organisation to the recipient.

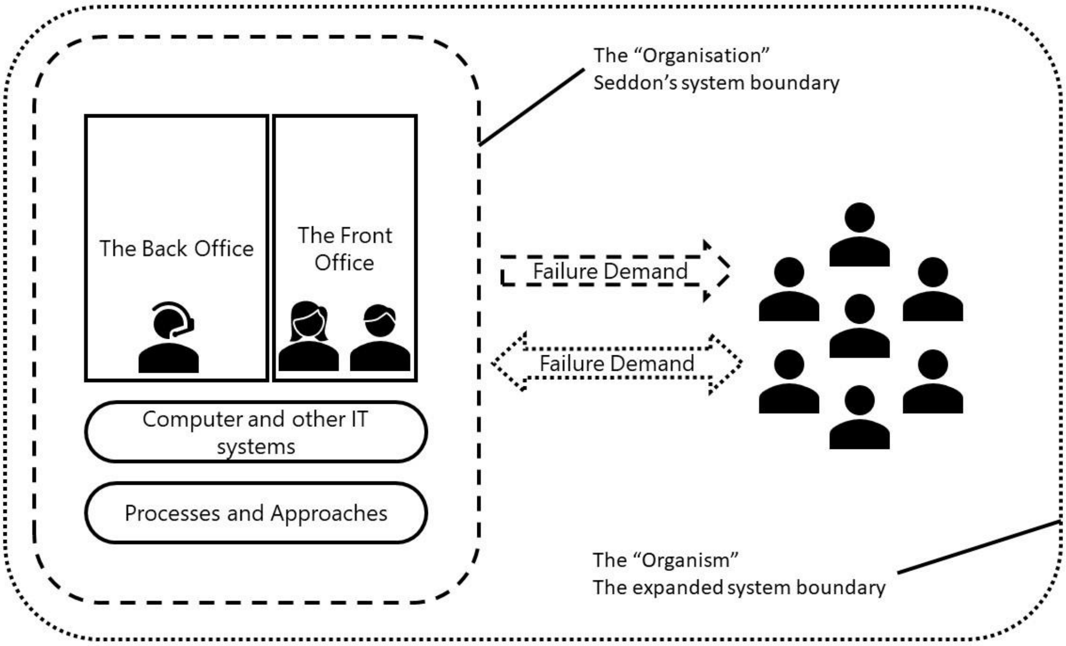


Figure 2-11: Our expanded system showing the organism rather than just the organisation.

An alternative system boundary "the organism" (the finely dotted line) is proposed, that draws the system domain to include the service recipient. This is necessary, as several failure demand events can be caused by the recipient. An appropriate anecdote would be a patient who (despite doctors' orders) eats prior to surgery and therefore causes failure demand. The actions, purposeful or accidental of service recipients may frequently lead to failure demand, and so, must be included as agents capable of its creation.

This modification of the system boundary also modifies the uni-directional arrow into a bi-directional arrow implying that as much as failure demand is in the control of the organisation to control, so too – failure demand in other instances is within the control of the service recipient.

In conversation with John Seddon – the importance of the idea of engineering the system to eliminate opportunities for the service recipient to cause failure demand was highlighted. As per the vanguard method [32], if failure demand occurrences are predictable, then this will imply that the mechanisms are systemic [32, 238]. If causes are systemic, then as per Deming [239], they must be eliminated. In many cases this is not easy.

2 - 9.2.3. **Why is failure demand problematic?**

The underlying reasons why failure demand is problematic affects both the organisation and the recipient. From the point of view of the organisation, resources that have been consumed to provide a service must be provided again. Conversely, from the point of view of the service recipient, looking from the outside in [32], failure demand causes inconvenience and frustration.

From the point of view of capacity, the inconvenience of recipients is not particularly problematic, however, when resources are consumed repetitively for the same purpose, this is wasteful, and as a detractor reduces effective system capacity [40], because resources are tied up doing the same work more than once. This means, that any work done prior to system failure may as well never have been done in the first place [240]. This "repeat consumption" is the underlying problem that is exposed when failure demand happens.

Failure demand matters because it decreases system capacity. It does so by repeatedly drawing on the same system resources or functions. In doing so – failure demand diminishes the ability of the system to deliver service to match demand. A common response to this, is that the system is designed with over-capacity to cope with demand which is higher than it needs to be leading to additional cost and system complexity.

2 - 9.2.4. **How failure demand is treated in this thesis**

Failure demand is difficult to measure because it manifests longitudinally. What is meant by this, is that failure demand only emerges when a demand instance returns to a system. In other words, if the system fails to deliver an adequate service to the recipient, and the consumer departs the system without having received resolution, this may be considered system failure, but has not yet compounded the depletion of the service capacity. Only when this unsuccessful instance of work returns in pursuit of the same unit of resolution at a later point, does this become failure demand.

The view expressed in this thesis is not pure “failure demand”. Key elements that drive this are the bidirectionality of agency for causing repeat consumption, the system boundary including recipient as equal agent in the production of failure and the focus and shifting towards repeat consumption. Rather than redefining failure demand on these nuances (as Masters did [240]) we clarify that for our purposes, the service recipient is within the system boundary, and therefore can be an agent in the emergence of failure demand. The thesis will equally treat failure demand as the mechanism by which a system failure is translated into the repeated consumption of system resources.

To explore this idea further, Chapter 5 provides an algorithm and a categorisation of commonly observed events that define system failure. A key idea in understanding system failure, lies in understanding "resolution" in other words, understanding whether a returning patient (say a chronic diabetic patient) is resolved or not, despite cure not having taken place. For this, the distinction between "care" and "cure", as proposed by Lillrank [241] is useful and as a general rule where cure is reasonably expected and not achieved, this may be system failure, whilst where only care is possible (say in dialysis), this does not represent system failure when such demand returns.

Failure demand is treated as a concept, that can be used to trouble shoot systems and identify behaviours which leads to an increase in demand. Its purpose is to highlight opportunities for improvement for which suitable interventions and tools should be designed [44, 110, 185]. Failure demand is not in itself a tool, but rather as a constituent of total demand, it represents a manageable portion of load on the system, which with suitable interventions can improve operations and reduce system load.

2 - 9.2.4.1. The curious case of NI 014

Although this thesis claims that failure demand in health care is an almost ignored element of assessing the quality of service-delivery, it is important to mention an instance where failure demand was used as a major reporting indicator, the eventual demise of this metric and its persistent bad reputation.

In 2007, The United Kingdom introduced a suite of national indicators for public services that are used to track key socio-economic indicators by definitively assessing national adherence to best practise policies and principles [242].

National Indicator 014, Alternatively referred to as NI 014 (or 14, without the leading zero [194]), tracked a phenomenon that was referred to as "avoidable contact". Scrutinising the definition that was provided by the British government [242] "avoidable contact", showed considerable similarity to the established definition for failure demand [44]. The definition that the British government used was:

NI 014 - Avoidable contact: the proportion of customer contact that is of low or no value to the customer the council failing to do something or do something right for the customer [243]

Avoidable Contacts would be measured as the total number of customer contacts against the total number of resolved customer requests expressed as a percentage [243]

John Seddon called the way that NI 014 was being implement "just plain wrong" [244 Pg. 1] and that the Cabinet Office simply "does not get it" [244 Pg. 5]. Seddon criticised that the implementation of this indicator was not approached from a systems point of view, in other words, not targeted towards the design of work and reducing a command and control culture. He bases this assessment on several arguments. He emphasised that failure demand is only a useful concept when it is used to identify systemic opportunities for improvement, and that this is only possible if the focus is on predictable failure demand. The nature of *predictable* failure demand is that it is *preventable.* But failure demand is only preventable if the systemic root can be found [44]. To do so failure demand should be treated as an indicator of the presence of systemic improvement opportunities [244]. But by being named "avoidable" rather than "preventable", the indicator became a target, rather than an analytical tool [244]. As a target, behaviours were automatically modified leading to managers applying "ingenuity [...] to meeting targets rather than [...] improving work" [244 Pg. 6]. This was further compounded by a general misunderstanding of the concept and large scale under-reporting [244].

The greatest difficulties that were reported by Walley et al., [194] included concerns derived from the LGITU magazine, which seems to now be defunct or is otherwise inaccessible. These issues could be distilled to four major impediments.

- The people were not ready to measure "avoidable contact",
- The technology was not set up properly to assist in understanding this phenomenon,
- The system was not integrated (including the service system and the IT system) and
- Analysis was difficult, expensive and there was poor data.

In addition to the environment that was not ready to measure "avoidable contact" there was no budget provided to make actual improvements, which remains the key purpose for measuring failure demand in the first place [44]. As a result, the project produced reporting of questionable quality, overloading administrators in doing so and was set up to deliver no systemic benefit.

Following on from difficulty in reporting. [194, 244] and errors and general controversy around its use [245] the metric was eventually quietly withdrawn from the list of national indicators in 2010 [194].

NI 014 was a metric deployed across government services and not specifically in health care only. Although the language, and the indicator were withdrawn quite soon from the National Indicators, it seems that elements of the idea have persisted. The NHS has advanced the avoidance of "avoidable consultations" [246]. The principles here were picked up to show how much of the demand in a practice derives from unnecessary contacts, and that managing these will re-establish efficiencies [246]. This interpretation of "avoidable consultations" focuses on diverting up to 27 % of consultations elsewhere, either to self-care, pharmacies or other health care providers [246]. McCarthy protests the implied ease with which this can be implemented, saying that "General practice is magnificently complicated, defying clean pathways and templates" [247 Pg. 1]. Here she unknowingly echoes Seddon's view that standardisation in a complex knowledge industry, particularly service and more specifically health care, may lead to increasing cost, rather than cost reduction.

Given its inappropriate implementation, and the difficulty that would have existed to overcome this, it is perhaps fortunate that NI 014 was removed as a national indicator. That should not serve to undermine the value of measuring failure demand in health care altogether, but that these interventions need to be done cautiously, correctly and for the right reasons. Failure demand should not be used as a target [244] but it should be used diagnostically [194]. Failure must be seen as a means for identifying systemic weaknesses and focus interventions for improvement [44] at the level of the root cause and not superficially as is common for managers who do not have a systems view [194].

2 - 10. References

3. Worldometer. *Government and Economics Indicators*. 2020 [cited 2020 15/03/2020].
8. Dijulio, B., Kirzinger, A., Wu, B., and Brodie, M., *Data Note: Americans' Challenges with Health Care Costs.* Kaiser Family Foundation, 2017.
9. Whitehurst, L., *Utah Sends Employees to Mexico for Lower Prescription Prices*, in *Associated Press News*. 2020, The Associated Press, New York.
13. Graban, M., *Lean Hospitals*. 2008, CRC Press, Boca Raton, London, New York
14. Ohno, T., *Toyota Production System: Beyond Large-Scale Production*. 1988, CRC Press, Boca Raton, London, New York
15. Liker, J.K., *The Toyota Way*. 2004, McGraw Hill, New York
19. Vickery, D.M. and Lynch, W.D., *Demand Management: Enabling Patients to Use Medical Care Appropriately.* Journal of Occupational and Environmental medicine, 1995. 37(5), p. 551-557.
21. Beall, J. *Beall's List* Beall's List of Predatory Journals and Publishers 2020 [cited 2020.
22. Isi Web of Science, *Isi Web of Science Database*. 2020, Clarivate Analytics, Philadelphia, USA.
23. Adams, R.J., Smart, P., and Huff, A.S., *Shades of Grey: Guidelines for Working with the Grey Literature in Systematic Reviews for Management and Organizational Studies.* International Journal of Management Reviews, 2017. 19(4), p. 432-454.
24. Conn, V.S., Valentine, J.C., Cooper, H.M., and Rantz, M.J., *Grey Literature in Meta-Analyses.* Nursing Research, 2003. 52(4), p. 256-261.
25. Mcauley, L., Tugwell, P., and Moher, D., *Does the Inclusion of Grey Literature Influence Estimates of Intervention Effectiveness Reported in Meta-Analyses?* The Lancet, 2000. 356(9237), p. 1228-1231.
26. Dickersin, K., Min, Y.-I., and Meinert, C.L., *Factors Influencing Publication of Research Results: Follow-up of Applications Submitted to Two Institutional Review Boards.* JAMA, 1992. 267(3), p. 374-378.
27. Dickersin, K. and Min, Y.-I., *Nih Clinical Trials and Publication Bias.* The Online Journal of Current Clinical Trials, 1993.

28. Easterbrook, P.J., Gopalan, R., Berlin, J., and Matthews, D.R., *Publication Bias in Clinical Research.* The Lancet, 1991. 337(8746), p. 867-872.
29. Rotonda, T., *Underreporting Research Is Scientific Misconduct.* JAMA, 1990. 263, p. 1405-1408.
30. Moraros, J., Lemstra, M., and Nwankwo, C., *Lean Interventions in Healthcare: Do They Actually Work? A Systematic Literature Review.* International Journal for Quality in Health Care, 2016. 28(2), p. 150-165.
31. Crotty, M., *The Foundations of Social Research: Meaning and Perspective in the Research Process*. 1998, Sage, New York City
32. Seddon, J., *The Vanguard Method.* Introduction to Vanguard - eLearning, 2020(1).
33. Seddon, J., *Failure Demand.* 2020.
34. Kingman, J.F.C., *The Single Server Queue in Heavy Traffic.* Mathematical Proceedings of the Cambridge Philosophical Society, 1961. 57(04), p. 902-904.
35. Kingman, J.F., *Two Similar Queues in Parallel.* The Annals of Mathematical Statistics, 1961. 32(4), p. 1314-1323.
36. Kingman, J.F.C., *On Queues in Heavy Traffic.* Journal of the Royal Statistical Society. Series B (Methodological), 1962. 24(2), p. 383-392.
37. Kingman, J.F.C., *On the Algebra of Queues.* Journal of Applied Probability, 1966. 3(2), p. 285-326.
38. Bicheno, J. *Towards Reducing Queues: Muri, Mura, Muda*. in *European Lean Educator Conference (ELEC2018)*. 2018. Braga, Portugal.
39. Bicheno, J., *Queues!*, D. Hartmann, Editor. 2019, Buckingham, United Kingdom.
40. Hopp, W.J. and Spearman, M.L., *Factory Physics*. 2011, Waveland Press, Long Grove, Illinois
41. Bicheno, J., *The King of Equations*, in *The Lean Management Journal*. 2011, SayOne Media, Norwich. p. 6.
42. Bicheno, J., *The Lean Games and Simulations Book*. 2015, Picsie Books, Buckingham, UK
43. Roser, C., *A Small Dice Game for the Kingman Formula*, in *All About Lean*, C. Roser, Editor. 2017.
44. Seddon, J., *Freedom from Command and Control: Rethinking Management for Lean Service*. 2005, CRC press, Boca Raton, Florida

45. Little, J.D. and Graves, S.C., *Little's Law*, in *Building Intuition*. 2008, Springer, New York City. p. 81-100.
46. Morse, P.M., *Queues, Inventories, and Maintenance*. 1958, Dover Publications Inc., Mineola, New York
47. Little, J.D., *A Proof for the Queuing Formula.* Operations Research, 1961. 9(3), p. 383-387.
48. Potter, A., Towill, D.R., and Gosling, J., *On the Versatility of Little's Law in Operations Management: A Review and Classification Using Vignettes.* Production Planning & Control, 2019, p. 1-16.
49. Jones, S.S., Thomas, A., Evans, R.S., Welch, S.J., Haug, P.J., and Snow, G.L., *Forecasting Daily Patient Volumes in the Emergency Department.* Academic Emergency Medicine, 2008. 15(2), p. 159-170.
50. Goldratt, E.M., *The Goal: A Process of Ongoing Improvement,* . 2004, Routledge, New York City
51. Brougham, G., *The Cynefin Mini-Book*. 2015, Lulu. com, Morrisville, North Carolina
52. Kurtz, C.F. and Snowden, D.J., *The New Dynamics of Strategy: Sense-Making in a Complex and Complicated World.* IBM Systems Journal, 2003. 42(3), p. 462-483.
53. Snowden, D., *Strategy in the Context of Uncertainty.* Handbook of Business strategy, 2005. 6(1), p. 47-54.
54. Snowden, D.J. and Boone, M.E., *A Leader's Framework for Decision Making.* Harvard Business Review, 2007. 85(11), p. 68.
55. Hasan, H. and Kazlauskas, A., *Making Sense of Is with the Cynefin Framework.* University of Wolongong, 2009.
56. Geron, S.M., *21st Century Strategies for Policing Protest: What Major Cities' Responses to the Occupy Movement Tell Us About the Future of Police Response to Public Protest*. 2014, Naval Postgraduate School, Monterey, California.
57. Snowden, D., *Journal of Knowledge Management.* Journal of knowledge management, 2002. 6(2), p. 100-111.
58. Lowy, A. and Hood, P., *The Power of the 2 X 2 Matrix: Using 2 X 2 Thinking to Solve Business Problems and Make Better Decisions*. 2011, John Wiley & Sons, Hoboken, New Jersey
59. Madsen, D.O., *Not Dead Yet: The Rise, Fall and Persistence of the Bcg Matrix.* Problems and Perspectives in Management, 2017. 15(1), p. 19-34.

60. Vollmar, J., Gepp, M., Palm, H., and Cala, A. *Engineering Framework for the Future: Cynefin for Engineers*. in *2017 IEEE International Systems Engineering Symposium (ISSE)*. 2017. IEEE.
61. Hasan, H. and Kazlauskas, A., *The Cynefin Framework: Putting Complexity into Perspective.* Univerity of Wolongong, 2014.
62. Snowden, D., *The Knowledge You Need, Right When You Need It.* Knowledge Management Review, 2003. 5, p. 24-27.
63. Bicheno, J. and Holweg, M., *The Lean Toolbox: A Handbook for Lean Transformations*. 5th Edition ed. 2016, PICSIE Books, Buckingham, UK
64. Jones, D.T. and Womack, J.P., *Lean Thinking*. 2012, Simon & Shuster, London
65. Womack, J.P., Jones, D.T., and Roos, D., *The Machine That Changed the World: How Lean Production Revolutionized the Global Car Wars*. 1991, Simon & Schuster, London
66. Krafcik, J.F., *Triumph of the Lean Production System.* MIT Sloan Management Review, 1988. 30(1), p. 41.
67. Lean Enterprise Institute. *A Brief History of Lean*. 2019 2019/04/02]; Available from: https://www.lean.org/WhatsLean/History.cfm.
68. Dolinsky, A., *Inventory Management History Part Three: Venetian Arsenal–Ahead of Their Time,*. 2013, Almyta Systems, Reno, Nevada.
69. Seddon, J., *Watch out for the Toolheads.* Lean-service. com [accessed 1 Feb 2005], 2005.
70. Smith, A. and Nicholson, J.S., *An Inquiry into the Nature and Causes of the Wealth of Nations*. 1887, T. Nelson and Sons, Edinburgh, UK
71. Whitney, E., *Correspondence of Eli Whitney Relative to the Invention of the Cotton Gin*. 2013, Read Books Ltd, Worcestershire, UK
72. Mccarthy, M. and Garcia, R., *Screw Threads on the Ss Xantho Engine: A Case of Standardisation in 19th Century Britain.* International Journal of Nautical Archaeology, 2004. 33(2), p. 330-337.
73. Woodbury, R.S., *The Legend of Eli Whitney and Interchangeable Parts.* Technology and Culture, 1960. 1(3), p. 235-253.
74. Gilbreth, F.B., *Motion Study: A Method for Increasing the Efficiency of the Workman*. 1911, D. Van Nostrand Company, New York City
75. Wren, D.A. and Bedeian, A.G., *The Evolution of Management Thought*. 1994, John Wiley & Sons, Hoboken, New Jersey

76. Fayol, H., *General and Industrial Management*. 2016, Ravenio Books, Cambridge, UK
77. Gilbreth, F.B., *Primer of Scientific Management*. 1912, D. Van Nostrand Co., New York
78. Taylor, F.W., *The Principles of Scientific Management*. 1914, Harper, New York
79. Gilbreth, F.B., *Bricklaying System*. 1909, The M.C. Clark publishing co., New York City
80. Gilbreth, F.B. and Gilbreth, L.M., *Fatigue Study, the Elimination of Humanity's Greatest Unnecessary Waste: A First Step in Motion Study*. 1916, Sturgis & Walton Company, New York City
81. Locke, E.A., *The Ideas of Frederick W. Taylor: An Evaluation.* Academy of Management Review, 1982. 7(1), p. 14-24.
82. Gilbreth, F.B. and Gilbreth, L.M., *Applied Motion Study a Collection of Papers on the Efficient Method to Industrial Preparedness.* 1917.
83. Khamara, E.J., *Space, Time, and Theology in the Leibniz-Newton Controversy*. Vol. 6. 2013, Walter de Gruyter
84. Cassirer, E., *Newton and Leibniz.* The Philosophical Review, 1943. 52(4), p. 366-391.
85. Taylor, F.W., *Scientific Management*. 2004, Routledge, Abingdon, UK
86. Ford, H., *Today and Tomorrow: Commemorative Edition of Ford's 1926 Classic*. 2018, Routledge, Abingdon, UK
87. Levinson, W.A., *Henry Ford's Lean Vision: Enduring Principles from the First Ford Motor Plant*. 2012, Productivity Press, New York City
88. Arnheiter, E.D. and Maleyeff, J., *The Integration of Lean Management and Six Sigma.* The TQM Magazine, 2005.
89. Mangaroo-Pillay, M., *A South African Ubuntu Analogy of the Lean Philosophy*, in *Department of Industrial Engineering, The North West University*. 2020, The North West University, Potchefstroom, South Africa.
90. Womack, J.P. and Jones, D.T., *Lean Thinking: Banish the Waste and Create in Your Corporation*. 1996, Simon and Schuster, London
91. Zidel, T.G., *A Lean Guide to Transforming Healthcare*. 2006, ASQ Quality Press, Milwaukee, Wisconsin
92. Williams, A., Kennedy, S., Philipp, F., and Whiteman, G., *Systems Thinking: A Review of Sustainability Management Research.* Journal of Cleaner Production, 2017. 148, p. 866-881.
93. Smuts, J.C., *Holism and Evolution.* 1926.

94. Becht, G., *Systems Theory, the Key to Holism and Reductionism.* Bioscience, 1974. 24(10), p. 569-579.
95. Checkland, P., *Systems Theory and Management Thinking.* American Behavioral Scientist, 1994. 38(1), p. 75-91.
96. Checkland, P., *Systems Thinking.* Rethinking Management Information Systems, 1999, p. 45-56.
97. Waltner-Toews, D., Kay, J., Kay, J.J., and Lister, N.-M.E., *The Ecosystem Approach: Complexity, Uncertainty, and Managing for Sustainability*. 2008, Columbia University Press
98. Bicheno, J., *The Service Systems Toolbox: Integrating Lean Thinking, Systems Thinking and Design Thinking*. 2012, Picsie Books, Buckingham, UK
99. Senge, P.M., *The Fifth Discipline.* Measuring Business Excellence, 1997. 1(3), p. 46-51.
100. Ackoff, R.L., *Management Misinformation Systems.* Management Science, 1967. 14(4), p. B-147-B-156.
101. Ackoff, R.L., Ackoff, R.L., and Emery, F.E., *On Purposeful Systems: An Interdisciplinary Analysis of Individual and Social Behavior as a System of Purposeful Events*. 2005, Transaction Publishers
102. Checkland, P., *From Optimizing to Learning: A Development of Systems Thinking for the 1990s.* The journal of the Operational Research Society, 1985. 36(9), p. 757-767.
103. Nakagawa, C.R.C., Nordholm, S., and Yan, W.-Y., *Analysis of Two Microphone Method for Feedback Cancellation.* IEEE Signal Processing Letters, 2014. 22(1), p. 35-39.
104. Homer, J.B. and Hirsch, G.B., *System Dynamics Modeling for Public Health: Background and Opportunities.* American Journal of Public Health, 2006. 96(3), p. 452-458.
105. Rusoja, E., Haynie, D., Sievers, J., Mustafee, N., Nelson, F., Reynolds, M., et al., *Thinking About Complexity in Health: A Systematic Review of the Key Systems Thinking and Complexity Ideas in Health.* Journal of Evaluation in Clinical Practice, 2018. 24(3), p. 600-606.
106. Lorenz, E., *The Butterfly Effect.* World Scientific Series on Nonlinear Science Series A, 2000. 39, p. 91-94.
107. Best, A. and Holmes, B., *Systems Thinking, Knowledge and Action: Towards Better Models and Methods.* Evidence & Policy: A Journal of Research, Debate and Practice, 2010. 6(2), p. 145-159.

108. Betzler, S.B., Wisnet, A., Breitbach, B., Mitterbauer, C., Weickert, J., Schmidt-Mende, L., et al., *Template-Free Synthesis of Novel, Highly-Ordered 3d Hierarchical Nb 3 O 7 (Oh) Superstructures with Semiconductive and Photoactive Properties.* Journal of Materials Chemistry A, 2014. 2(30), p. 12005-12013.
109. Luhmann, N., *Social Systems*. 1995, stanford university Press
110. Piercy, N. and Rich, N., *Lean Transformation in the Pure Service Environment: The Case of the Call Service Centre.* International Journal of Operations & Production Management, 2009. 29(1), p. 54-76.
111. Gupta, S. and Sharma, M., *Lean Services: A Systematic Review.* International Journal of productivity and performance management, 2016.
112. Sasser, W.E., *Match Supply and Demand in Service Industries.* Harvard Business Review, 1976. 54(6), p. 133-140.
113. Levitt, T., *Production-Line Approach to Service.* Harvard Business Review, 1972. 50(5), p. 41-52.
114. Levitt, T., *Industrialization of Service.* Harvard Business Review, 1976. 54(5), p. 63-74.
115. Radnor, Z. and Walley, P., *Learning to Walk before We Try to Run: Adapting Lean for the Public Sector.* Public Money and Management, 2008. 28(1), p. 13-20.
116. Radnor, Z. and Osborne, S.P., *Lean: A Failed Theory for Public Services?* Public Management Review, 2013. 15(2), p. 265-287.
117. Bowen, D.E. and Youngdahl, W.E., *"Lean" Service: In Defense of a Production-Line Approach.* International Journal of Service Industry Management, 1998.
118. Spear, S.J., *Fixing Health Care from the inside, Today.* Harvard Business Review, 2005. 83(9), p. 78.
119. Seddon, J., *Watch out for the Toolheads! Everything You Need to Know About Lean Manufacturing Tools and Why They Won't Work in Service Organisations*. 2005, Systemthinkingmethod.com.
120. Radnor, Z., Walley, P., Stephens, A., and Bucci, G., *Evaluation of the Lean Approach to Business Management and Its Use in the Public Sector.* Scottish Executive Social Research, 2006. 20, p. 1-6.
121. Jones, D.T. and Womack, J.P., *Lean Consumption.* Harvard Business Review, 2005. 83, p. 58-68.
122. Womack, J., *Lean Consumption.* Manufacturing Engineer, 2005. 84(4), p. 5-5.

123. Allway, M. and Corbett, S., *Shifting to Lean Service: Stealing a Page from Manufacturers' Playbooks.* Journal of Organizational Excellence, 2002. 21(2), p. 45-54.
124. Oed Online, *The Oxford English Dictionary*, in *The Oxford English Dictionary*. 2020, Oxford University Press, Oxford, UK.
125. Cobra, M., *Serviços: Como Contruir Valor Para O Cliente*. 2004, Marcos Cobra Editora Ltda
126. Leite, H.D.R. and Vieira, G.E., *Lean Philosophy and Its Applications in the Service Industry: A Review of the Current Knowledge.* Production, 2015. 25(3), p. 529-541.
127. George, M., *Lean Six Sigma for Service*. 2003, McGraw Hill Professional, New York City
128. Johnston, R., *Service Operations Management: Return to Roots.* International Journal of Operations & Production Management, 1999. 19(2), p. 104 - 124.
129. Andrés-López, E., González-Requena, I., and Sanz-Lobera, A., *Lean Service: Reassessment of Lean Manufacturing for Service Activities.* Procedia Engineering, 2015. 132(1), p. 23-30.
130. Roy, R.B., Lillrank, P., Sreekanth, V., and Torkki, P., *The Conceptual Tools*, in *Designing Service Machines*. 2019, Springer. p. 5-19.
131. Roy, R.K., *A Primer on the Taguchi Method*. 2010, Society of Manufacturing Engineers, Ontario, Canada
132. Teece, D.J., *Business Models, Business Strategy and Innovation.* Long Range Planning, 2010. 43(2-3), p. 172-194.
133. Schneider, D. and Schnetkamp, G., *E-Markets: B2b-Strategien Im Electronic Commerce, Marktplätze, Fachportale, Plattformen*. 2013, Springer-Verlag, Wiesbaden
134. Seddon, J. and Brand, C., *Debate: Systems Thinking and Public Sector Performance.* Public Money & Management, 2008. 28(1), p. 7-9.
135. Skinner, W., *Manufacturing-Missing Link in Corporate Strategy.* Harvard Business Review, 1969, p. 136 - 145.
136. Seddon, J., *I Want You to Cheat!: The Unreasonable Guide to Service and Quality in Organisation*. 1 ed. 1992, Vanguard Press, Buckingham, UK
137. Seddon, J., O'donovan, B., and Zokaei, K., *Rethinking Lean Service*, in *Service Design and Delivery*. 2011, Springer. p. 41-60.
138. Seddon, J., *Lean Is a Waning Fad.* Management Services, 2011. 55(4), p. 34-36.

139. Radnor, Z., *Transferring Lean into Government.* Journal of Manufacturing Technology Management, 2010.
140. Lawal, A.K., Rotter, T., Kinsman, L., Sari, N., Harrison, L., Jeffery, C., et al., *Lean Management in Health Care: Definition, Concepts, Methodology and Effects Reported (Systematic Review Protocol).* Systematic reviews, 2014. 3(1), p. 1-6.
141. Radnor, Z.J., Holweg, M., and Waring, J., *Lean in Healthcare: The Unfilled Promise?* Social Science & Medicine, 2012. 74(3), p. 364-371.
142. Antony, J., Sreedharan, R., Chakraborty, A., and Gunasekaran, A., *A Systematic Review of Lean in Healthcare: A Global Prospective.* International Journal of Quality & Reliability Management, 2019.
143. Preston, N.J., *The Health of the Nation Outcome Scales: Validating Factorial Structure and Invariance across Two Health Services.* Australian & New Zealand Journal of Psychiatry, 2000. 34(3), p. 512-519.
144. Mason, S., Nicolay, C., and Darzi, A., *The Use of Lean and Six Sigma Methodologies in Surgery: A Systematic Review.* The Surgeon, 2015. 13(2), p. 91-100.
145. Burgess, N. and Radnor, Z., *Evaluating Lean in Healthcare.* International Journal of Health Care Quality Assurance, 2013. 26(3), p. 220 - 235.
146. Weimarsson, H., *Nine out of Ten Hospitals Changed to" Lean".* Lakartidningen, 2011. 108(39), p. 1915-1915.
147. Institutet, K. *Website of the Karolinska Institutet*. 2020 [cited 2020 2020]; Available from: https://ki.se/en.
148. Ker, J.-I., Wang, Y., Hajli, M.N., Song, J., and Ker, C.W., *Deploying Lean in Healthcare: Evaluating Information Technology Effectiveness in Us Hospital Pharmacies.* International Journal of Information Management, 2014. 34(4), p. 556-560.
149. Hartmann, D.H. *Pharmacy Optimisation: Getting People on Drugs Faster*. in *International Conference on Industrial Engineering, Systems Engineering and Engineering Management*. 2011. Stellenbosch.
150. Kachwee, M. and Hartmann, D. *Hospital Supply Chain Management and Optimisation*. in *Southern African Institute for Industrial Engineering (SAIIE) 25*. 2013. Stellenbosch.
151. Ryan, A., Hunter, K., Cunningham, K., Williams, J., O'shea, H., Rooney, P., et al., *Steps: Lean Thinking, Theory of Constraints and*

Identifying Bottlenecks in an Emergency Department. Irish Medical Journal, 2013. 106(4), p. 105-7.
152. Hartmann, D. and Mandavha, R. *Lean Healthcare, a Casualty of Inefficiency.* in *SAIIE Conference*. 2010. Pretoria, Southern African Institute for Industrial Engineering.
153. Nicolay, C., Purkayastha, S., Greenhalgh, A., Benn, J., Chaturvedi, S., Phillips, N., et al., *Systematic Review of the Application of Quality Improvement Methodologies from the Manufacturing Industry to Surgical Healthcare*, in *Database of Abstracts of Reviews of Effects (Dare): Quality-Assessed Reviews [Internet]*. 2012, Centre for Reviews and Dissemination (UK).
154. Hartmann, D. and Sunjka, B., *Private Theatre Utilisation in South Africa: A Case Study.* South African Medical Journal, 2013. 103(5), p. 285-7.
155. Hartmann, D. and Sooklal, S., *The Pen Is Mightier Than the Scalpel : The Case for Electronic Medical Records : Case Study.* South African Journal of Industrial Engineering, 2012. 23(2), p. 191 - 201.
156. Silvester, K., Lendon, R., Bevan, H., Steyn, R., and Walley, P., *Reducing Waiting Times in the Nhs: Is Lack of Capacity the Problem?* Clinician in Management, 2004. 12(3), p. 105-109.
157. Kenney, C., *Transforming Health Care: Virginia Mason Medical Center's Pursuit of the Perfect Patient Experience*. 2012, CRC Press, Boca Raton, Florida
158. Weber, D.O., *Toyota-Style Management Drives Virginia Mason.* Physician Executive, 2006. 32(1), p. 12.
159. Meredith, J.O., Grove, A.L., Walley, P., Young, F., and Macintyre, M.B., *Are We Operating Effectively? A Lean Analysis of Operating Theatre Changeovers.* Operations Management Research, 2011. 4(3-4), p. 89.
160. Scherrer-Rathje, M., Boyle, T.A., and Deflorin, P., *Lean, Take Two! Reflections from the Second Attempt at Lean Implementation.* Business horizons, 2009. 52(1), p. 79-88.
161. Mazzocato, P., Savage, C., Brommels, M., Aronsson, H., and Thor, J., *Lean Thinking in Healthcare: A Realist Review of the Literature.* BMJ Quality & Safety, 2010. 19(5), p. 376-382.
162. Lucey, J., Bateman, N., and Hines, P., *Why Major Lean Transitions Have Not Been Sustained.* Management Services, 2005. 49(2), p. 9-13.

163. Andersen, H., Røvik, K.A., and Ingebrigtsen, T., *Lean Thinking in Hospitals: Is There a Cure for the Absence of Evidence? A Systematic Review of Reviews.* BMJ open, 2014. 4(1).
164. Glasgow, J.M., Scott-Caziewell, J.R., and Kaboli, P.J., *Guiding Inpatient Quality Improvement: A Systematic Review of Lean and Six Sigma.* The Joint Commission Journal on Quality and Patient Safety, 2010. 36(12), p. 533-AP5.
165. Brandão De Souza, L., *Trends and Approaches in Lean Healthcare.* Leadership in Health Services, 2009. 22(2), p. 121-139.
166. Alexander, J.A. and Hearld, L.R., *What Can We Learn from Quality Improvement Research? A Critical Review of Research Methods.* Medical Care Research and Review, 2009. 66(3), p. 235-271.
167. Radnor, Z. and Boaden, R., *Editorial: Lean in Public Services—Panacea or Paradox?* Public Money & Management, 2008. 28(1), p. 3 - 7.
168. Leseure, M., Hudson-Smith, M., and Radnor, Z., *Transferring Lean into Government.* Journal of Manufacturing Technology Management, 2010.
169. Kim, C.S., Spahlinger, D.A., Kin, J.M., and Billi, J.E., *Lean Health Care: What Can Hospitals Learn from a World-Class Automaker?* Journal of Hospital Medicine, 2006. 1(3), p. 191-199.
170. Jones, D. and Mitchell, A., *Lean Thinking for the Nhs.* London: NHS Confederation, 2006.
171. Rotter, T., Plishka, C., Lawal, A., Harrison, L., Sari, N., Goodridge, D., et al., *What Is Lean Management in Health Care? Development of an Operational Definition for a Cochrane Systematic Review.* Evaluation & the Health Professions, 2019. 42(3), p. 366-390.
172. Sloan, T., Fitzgerald, A., Hayes, K.J., Radnor, Z., Robinson, S., Sohal, A., et al., *Complexity Complicates Lean: Lessons from Seven Emergency Services.* Journal of Health Organization and Management, 2014. 28(2), p. 266 - 288.
173. Fujimoto, T., *The Evolution of a Manufacturing System at Toyota*. 1999, Oxford University Press, Oxford, UK
174. Hines, P., Holweg, M., and Rich, N., *Learning to Evolve: A Review of Contemporary Lean Thinking.* International Journal of Operations & Production Management, 2004. 24(10), p. 994-1011.
175. Mazzocato, P., Holden, R.J., Brommels, M., Aronsson, H., Bäckman, U., Elg, M., et al., *How Does Lean Work in Emergency Care? A Case Study of a Lean-Inspired Intervention at the Astrid Lindgren*

Children's Hospital, Stockholm, Sweden. BMC Health Services Research, 2012. 12(1), p. 1-13.
176. Plsek, P.E. and Greenhalgh, T., *Complexity Science: The Challenge of Complexity in Health Care.* BMJ: British Medical Journal, 2001. 323(7313), p. 625.
177. Plsek, P.E. and Wilson, T., *Complexity, Leadership, and Management in Healthcare Organisations.* BMJ: British Medical Journal, 2001. 323(7315), p. 746.
178. Burgess, N. and Radnor, Z., *Thesis: Evaluating Lean in Healthcare*. 2012, University of Warwick, Warwick, UK.
179. Ingebrigtsen, T., *Merging of Three Hospitals into One University Hospital 813–7.*
180. Proudlove, N., Moxham, C., and Boaden, R., *Lessons for Lean in Healthcare from Using Six Sigma in the Nhs.* Public Money & Management, 2008. 28(1), p. 27-34.
181. Trist, E., *The Evolution of Socio-Technical Systems.* Occasional Paper, 1981. 2, p. 1981.
182. Walshe, K., *Understanding What Works—and Why—in Quality Improvement: The Need for Theory-Driven Evaluation.* International Journal for Quality in Health Care, 2007. 19(2), p. 57-59.
183. Mcnulty, T. and Ferlie, E., *Reengineering Health Care: The Complexities of Organizational Transformation*. 2002, OUP Oxford, Oxford, UK
184. Seddon, J., *Discussion of Failure Demand Repeat Consumption and the Idea of Demand*, D. Hartmann, Editor. 2020, Johannesburg, South Africa.
185. Marr, B. and Neely, A., *Managing and Measuring for Value: The Case of Call Centre Performance*. 2004, Cranfield School of Management and Fujitsu, Cranfield, UK.
186. Walley, P., *Does the Public Sector Need a More Demand-Driven Approach to Capacity Management?* Production Planning & Control, 2013. 24(10-11), p. 877-890.
187. O'sullivan, A. and Sheffrin, S.M., *Economics: Principles in Action*. 2007, Pearson/Prentice Hall, Upper Saddle River, New Jersey
188. Mankiw, N.G., *Principles of Economics*. 2014, Nelson Education, Toronto, Canada
189. Demir, E., Gunal, M.M., and Southern, D., *Demand and Capacity Modelling for Acute Services Using Discrete Event Simulation.* Health Systems, 2017. 6(1), p. 33-40.

190. Allder, S., Silvester, K., and Walley, P., *Managing Capacity and Demand across the Patient Journey.* 2010. 10(1), p. 13.
191. Lovelock, C.H., *Strategies for Managing Demand in Capacity-Constrained Service Organisations.* The Service Industries Journal, 1984. 4(3), p. 12-30.
192. Allder, S., Walley, P., and Silvester, K., *Is Follow-up Capacity the Current Nhs Bottleneck?* Clinical Medicine, 2011. 11(1), p. 31.
193. Walley, P. and Adams, M.M., *An Evaluation of Demand Management Practices in Uk Police Forces.* 2019.
194. Walley, P., Found, P., and Williams, S., *Failure Demand: A Concept Evaluation in Uk Primary Care.* International Journal of Health Care Quality Assurance, 2019. 32(1), p. 21-33.
195. Praeter, J., *Queues in Health.* Health Care Management Science, 2002. 5(4), p. 283 - 283.
196. Hartmann, D. *I Just Don't Have Enough Time*. in *Southern African Institute for Industrial Engineering*. 2010. Johannesburg.
197. Toussaint, J. and Gerard, R., *On the Mend: Revolutionizing Healthcare to Save Lives and Transform the Industry*. 2010, Lean Enterprise Institute, Cambridge, Massachusetts
198. Wagstaff, A., Van Doorslaer, E., and Paci, P., *On the Measurement of Horizontal Inequity in the Delivery of Health Care.* Journal of Health Economics, 1991. 10, p. 169-205.
199. Crandall, R.E. and Markland, R.E., *Demand Management-Today's Challenge for Service Industries.* Production and Operations Management, 1996. 5(2), p. 106-120.
200. Hosseini, H., *Understanding the Market Mechanism before Adam Smith: Economic Thought in Medieval Islam.* History of Political Economy, 1995. 27(3), p. 539-561.
201. Marshall, A., *Principles of Economics*. 2004, Digireads. com Publishing
202. Macario, A., *Are Your Operating Theatres Being Run Efficiently?* Medscape Anesthesiology, 2010.
203. Schoen, C., Osborn, R., Doty, M.M., Bishop, M., Peugh, J., and Murukutla, N., *Toward Higher-Performance Health Systems: Adults' Health Care Experiences in Seven Countries, 2007.* Health Aff (Millwood), 2007. 26(6), p. w717-34.
204. South African National Department of Health, *National Department of Health Strategic Plan 2010/2011 - 2012/2013*, S.A.N.D.o. Health, Editor. 2010, Pretoria. p. 115.

205. South African National Department of Health, *Strategic Plan 2014/15 - 2018/19*. 2014, South African National Department of Health,, Pretoria, South Africa.
206. Louis M. Savary, C.C.-M., *The Nun and the Bureaucrat: How They Found an Unlikely Cure for America's Sick Hospitals*. 2006, CC-M Productions, Washington
207. Hartmann, D., *Improvement in Operating Theatre Efficiency through Better Measurement and Scheduling Msc. Research Report*. 2012, University of the Witwatersrand Press, Johannesburg.
208. Heskett, J.L., *Service Breakthroughs*. 1990, Simon and Schuster, New York City
209. Klassen, K.J. and Rohleder, T.R., *Combining Operations and Marketing to Manage Capacity and Demand in Services.* Service Industries Journal, 2001. 21(2), p. 1-30.
210. Jack, E.P. and Powers, T.L., *A Review and Synthesis of Demand Management, Capacity Management and Performance in Health-Care Services.* International Journal of Management Reviews, 2009. 11(2), p. 149-174.
211. Hopp, W.J. and Lovejoy, W.S., *Hospital Operations: Principles of High Efficiency Health Care*. 2012, FT Press, Upper Saddle River, New Jersey
212. Naidoo, J. and Wills, J., *Health Promotion: Foundations for Practice*. 2010, Baillière Tindall, Sidcup, UK
213. Hogg, W., Rowan, M., Russell, G., Geneau, R., and Muldoon, L., *Framework for Primary Care Organizations: The Importance of a Structural Domain.* International Journal for Quality in Health Care, 2008. 20(5), p. 308-313.
214. World Health Organization, *Bangladesh Health System Review*. 2015, Manila: WHO Regional Office for the Western Pacific
215. Maeda, A., Araujo, E., Cashin, C., Harris, J., Ikegami, N., and Reich, M.R., *Universal Health Coverage for Inclusive and Sustainable Development: A Synthesis of 11 Country Case Studies*. 2014, The World Bank, Washington D.C.
216. El-Saharty, S., Sparkes, S.P., Barroy, H., Ahsan, K.Z., and Ahmed, S.M., *The Path to Universal Health Coverage in Bangladesh: Bridging the Gap of Human Resources for Health*. 2015, The World Bank, Washington D.C.

217. Unctad. *Un List of Least Developed Countries*. 2020 [cited 2020; Available from: https://unctad.org/topic/vulnerable-economies/least-developed-countries/list.
218. Osborn, R., Squires, D., Doty, M.M., Sarnak, D.O., and Schneider, E.C., *In New Survey of Eleven Countries, Us Adults Still Struggle with Access to and Affordability of Health Care.* Health Affairs, 2016. 35(12), p. 2327-2336.
219. Richardson, J., *Supply and Demand for Medical Care: Or, Is the Health Care Market Perverse?* Australian Economic Review, 2001. 34(3), p. 336-352.
220. Tajeu, G., *Health Maintenance Organization (Hmo)*, in *The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society*. 2014, John Wiley & Sons, Ltd, Hoboken, New Jersey.
221. Heineke, J., *Strategic Operations Management Decisions and Professional Performance in Us Hmos.* Journal of Operations Management, 1995. 13(4), p. 255-272.
222. Jack, E.P. and Powers, T.L., *Volume Flexible Strategies in Health Services: A Research Framework.* Production and Operations Management, 2004. 13(3), p. 230-244.
223. Lewin, M.E. and Altman, S., *America's Health Care Safety Net: Intact but Endangered*. 2000, National Academies Press, Washington D.C.
224. World Health Organization. *Naming the Coronavirus Disease (Covid-19) and the Virus That Causes It*. 2020; Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it.
225. The Straits Times, *Coronavirus: Wuhan to Build 19 More Makeshift Hospitals as China Struggles to Contain Epidemic*, in *The straits times*. 2020, Singapore Press Holdings, Singapore, Singapore.
226. World Health Organization, *Who Director-General's Opening Remarks at the Media Briefing on Covid-19 - 11 March 2020*. 2020, World Health Organization, .
227. Schuler, R., *60 Bis 70 Prozent in Deutschland Werden Sich Infizieren*, in *Bild*. 2020, Axel Springer SE, Berlin, Germany.
228. Regalado, A., *Worst-Case Coronavirus Scenario: 214 Million Americans Infected, 1.7 Million Dead*, in *MIT Technology Review*. 2020, MIT Technology Review Cambridge Massachusetts.
229. Qualls, N., Levitt, A., Kanade, N., Wright-Jegede, N., Dopson, S., Biggerstaff, M., et al., *Community Mitigation Guidelines to Prevent*

Pandemic Influenza—United States, 2017. MMWR Recommendations and Reports, 2017. 66(1), p. 1.
230. World Health Organization. *Coronavirus Disease (Covid-19) Advice for the Public*. 2020 [cited 2020; Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public.
231. Lorig, K.R., Mazonson, P.D., and Holman, H.R., *Evidence Suggesting That Health Education for Self-Management in Patients with Chronic Arthritis Has Sustained Health Benefits While Reducing Health Care Costs.* Arthritis & Rheumatology, 1993. 36(4), p. 439-446.
232. Donnelley, R., *Mental Health Collaborative Programme Toolkit*, T.S. Government, Editor. 2009, Crown, Edinburgh.
233. Healthcare Improvement Scotland, *A Guide to Understanding and Managing Demand*. 2019, Mental Health Access Improvement Support Team (MHAIST), Edinburgh, UK.
234. El-Ghorr, A., Cameron, R., Fleming, M., Mckechnie, L., Thomson, D., and Doherty, S., *Scotland's National Approach to Improving Mental Health Services: Integrated Care Pathways as Tools for Redesign and Continuous Quality Improvement.* 2010. 14(2), p. 57-64.
235. Seddon, J. and Brand, C., *Debate: Systems Thinking and Public Sector Performance.* 2008.
236. Donaldson, C. and Gerard, K., *Paying General Practitioners: Shedding Light on the Review of Health Services.* Journal of the Royal College of General Practitioners, 1989. 39(320), p. 114-117.
237. Richardson, J.R. and Peacock, S.J., *Supplier-Induced Demand.* Applied Health Economics and Health Policy, 2006. 5(2), p. 87-98.
238. United States Department of Health Human Services, *Healthy People 2000: National Health Promotion and Disease Prevention Objectives*, in *Healthy People 2000: National Health Promotion and Disease Prevention Objectives*. 1991, US Government Printing Office.
239. Deming, E., *Out of the Crisis. Quality, Productivity and Competitive Position. Cambridge, Ma: Center for Advanced Engineering Study*. 1986, Cambridge University Press.
240. Masters, K.I.A., *From New Public Management to Lean Thinking: Understanding and Managing'potentially Avoidable Failure Induced Demand'*. 2010, University of Sussex.

241. Lillrank, P., Groop, P.J., and Malmström, T.J., *Demand and Supply–Based Operating Modes—a Framework for Analyzing Health Care Service Production.* Milbank Quarterly, 2010. 88(4), p. 595-615.
242. Hm Government, *National Indicators for Local Authorities and Local Authority Partnerships: Handbook of Definitions*, C. Office, Editor. 2007, DCLG, London, UK.
243. Hm Government, *Ni 014 - Avoidable Contact:*. 2010, Cabinet Office, London, UK.
244. Seddon, J., *Ni 14 ('Avoidable Contact'):–from the Horse's Mouth'.* Customer Strategy, 2009. 2(1), p. 33-34.
245. Martin, P., *'Avoidable Contact: If Only...*. 2017, Local Government Chronicle.
246. Clay, H. and Stern, R., *Making Time in General Practice.* Birmingham: NHS Alliance, 2015, p. 59-63.
247. Mccartney, M., *Pointless Paperwork, Not Patients, Is What Gps Should Avoid.* BMJ : British Medical Journal (Online), 2017. 359.

Chapter Three: **Methods and Approach**

3 - 1. Introduction

This chapter first establishes the philosophical position of this work. It discusses the way in which this work was positioned in terms of its ontology and its epistemology which leads to the conclusion that the overall guiding philosophy of this thesis is one of positivism. This follows from the fact that demand can be observed in the real world independent of constructs in the human mind, can be measured and assessed and tested against collectible data.

Once this position has been established, this section describes several methods that were chosen, yet later abandoned due to a variety of reasons. This shows the evolution of how access to information changed as the complexities of the purpose of this thesis emerged.

Following on from that, a traditional method section is presented, that shows the overall structure of the study, where the individual questions are addressed and what data sources are used for each of the phases.

The chapter describes the two overarching phases of this thesis, the Conceptual phase, where the frameworks are built (Chapters 4 and 5 and then the Evaluative phase where an empirical study is performed to assess failure demand levels in a pharmaceutical supply-chain (Chapter 6). In each of these, the major analytical and other methods are described as are data sources and approaches towards validation.

The chapter concludes with a note on the way in which the Human Ethics Research Council indirectly affected the scope of the study due to previous experiences of access being limited to vulnerable populations, which largely includes patients.

3 - 2. "Beliefs Truth and Knowledge" [248]

Research is broadly split into two polarised groups [249]. The so-called "social sciences" and the "natural sciences" [16, 31]. These poles broadly approach work from different philosophical departure points and so make use of different methodologies. Social scientists are more likely to be qualitative researchers, whilst the natural sciences favour quantitative approaches [250]. This division manifests in the polarisation of world views and the purpose of research [251]. Social sciences strive first to understand the world, and to describe the complex fabric of humankind by using textured techniques to describe and understand the world [16]. By contrast natural scientists are more likely to seek data and other means of support to test theories and explain phenomena [251].

The idea of questioning beliefs as truth, and truth as knowledge is perhaps not a standard practise in the natural sciences. By comparison, the social sciences take great care to express the implicit and underlying philosophies and beliefs that frame research [252]. The social sciences carefully question what constitutes truth, and how beliefs are woven into the fabric of theory [31]. This introduces two important ideas, Ontology and Epistemology.

Ontology deals with the question of what reality is, in other words, ontology aims to answer the question of how we understand existence [251, 253].

Epistemology is the means by which humans assess whether knowledge is valid, proposes how it should be obtained, processed and trusted [251, 253]. The remainder of this section will explore the notions of ontology and epistemology. Once the supporting theory has been presented, this section will develop the philosophical position of the researcher in this thesis.

3 - 2.1. Ontologies

Realism proposes an observable world in which facts are divorced from the researcher's beliefs [254, 255]. Relativism on the other hand views truth as an intrinsically social device that cannot be divorced from the perspectives, feelings, opinions and biases of the agents who play a role in its construction [31, 254, 255].

3 - 2.1.1. **Realism**

The foundation of the realist world view is the idea that truth exists, and that it is not obscured, or variable, based on the mores or opinions of society [251, 254, 255]. Al Saadi's meta-analysis of sources shows that realism is characterised by [254]:

- Truth exists and is not shaped by beliefs,
- truth is directly observable and can be accurately measured,
- the metaphysical world is not considered real,
- causality exists and can be observed, interpreted and sometimes determined,
- truth emerges from measurement rather than experience and belief,
- social phenomena and their meanings are static, and
- "choice, freedom, individuality and moral responsibility are excluded" [254 Pg. 4].

3 - 2.1.2. **Relativism**

Relativism is characterised by the notion that meaning and existence are constructions based on human observation, knowledge and to an extent fashion [31, 255-257]. The nature of reality is dynamic and as Shakespeare expresses in Sonnet 116, "alters as it alteration finds". Thus, reality is relative to the observer, the context and the vicissitudes of life.

Al Saadi's [254] work, read together with Moon [256] and Crotty [31], shows that relativism is characterised by:

- Reality does exist but it is a human construct,
- no universal social truth exists, merely a variety of individual truths,
- truth can only be understood as an estimate or in approximate terms,
- social interaction is the engine that drives and creates meaning, which remains dynamic and ever-changing.

3 - 2.2. Epistemologies

3 - 2.2.1. **Objectivism**

In general, engineering is a realist discipline [258]. an ontology in which objective epistemologies are prevalent [249]. In this "facts are facts" universe [259] meaning can be "discovered where it has existed all along" [254 Pg. 3] It is easy to make the mental shift to the caricature of a scientist in a lab, who explores, measures and has no need to question the fundamental truth of measurement, readings and values [260].

On a deeper level, the objective approach is characterised by the need to explain phenomena, by use of objective knowledge which is constructed by deductive means [31]. This means that generally ideas are tested producing objective findings [16, 255]. These findings are characterised by originating from data, with an emphasis on measurement, coordination and statistical analysis [31, 260]. Ultimately, objectivist research aims to produce universal principles, which should ultimately result in generalisable findings [251, 258].

3 - 2.2.2. **Interpretivism**

The Interpretative epistemology differs from the objectivist in many key respects. Rather than a scientist, the researcher is now a detective [251], and seeks to understand, rather than to explain [260]. Knowledge is subjective and measured for validity through a highly contextualised lens [260]. Rather than a focus on facts, interpretative work focuses on "people being people" and the emergent construction of a perception of truth [251, 260]. An interpretivist seeks to understand how and why things happen and builds theories from observations and other sources. As Goertz calls it, understanding the substance of the empirical world [257].

3 - 2.3. The positivist philosophical position

A realist - objectivist axis is often a positive research philosophy - a world in which facts are facts [31, 255, 256]. Positivist work exists in a granular world that is real, independent, ordered and universal. The basis of positivist science is measurable or observable information, striving for a generalisable output [31, 255]. The quest for causality is central in positivism [31, 255] even if correlations are more frequently presented. Positivism aims to perform research in the absence of a value system which could drive or constrain the truth. This is achieved through a researcher who "keeps their distance" establishing an objective stance [255]. Leading on from the objectivity, the methods are usually deductive and structured [31, 255]. Measurements are taken, large amounts of data collected and numerical techniques are used to produce outputs [255].

3 - 2.4. The philosophical position of this thesis

This thesis investigates demand in health care and its manifestation in the real world. It is the position of the thesis, that demand remains real and present, regardless of whether it is observed, assessed, or has value systems applied to it.

Equally, failure demand manifests regardless of whether it is observed, and perhaps one of the important contributions of this thesis is that failure demand exists precisely whilst frequently *not* being observed (as do other forms of demand). Similarly, value systems do not modify the presence of demand – whether demand is "good" or "bad" does not alter the fact that demand exists and can positively be measured, analysed and described based on data collected about the behaviour of the health care system. This is characteristic of a positivist philosophical position – truth is distinct from human consciousness [253] and cause and effect relationships are presumed to exist and repeated observation allows their description.

The pilot studies that are described in Section 1 - 1.1 and Section 3 - 3 reveal that there were gaps in concepts, particularly related to a structured approach for understanding what failure demand in health care is, and how demand in health systems is structured. To build these theoretical models

[258], the researcher had to make use of some Interpretivist approaches, to construct the mental models that have been presented in Chapters 4 and 5 After the models had been constructed, they were tested against common scenarios thus again grounding the study in a positivist paradigm. Ahmed comments on this kind of study and views the quantitative testing of frameworks as overall, positivist [261].

A simple self-test derived from Raddon et al. is to ask whether, over the totality of the thesis, the researcher was a detective or a scientist [251]. It is the author's view that he was more scientist than detective. That the claims in the thesis are based on analytical findings grounded in theoretical models and data, with analysis (especially in Chapter 6) being empirical in nature and that the key output of this thesis is not based on biases, value systems, human perception or interpretation.

From this examination of the dominant modes of ontology and epistemology, and on to the ultimate purpose of this thesis, this study is framed in a realist ontology, where demand exists in the real world where meaning can exist outside of the human mind [31, 256]. Because knowledge is validated through testing and experimentation, this study was mostly conducted with an objectivist epistemology. This combination describes the philosophical position of positivism, the approach to science, knowledge and discovery.

The methods that were used in this thesis fell broadly into the positivist realm. But not all were used in this study. The following section briefly discusses which methods had to be abandoned, and why.

3 - 3. Abandoned Methods

The initial objective for this study was to get a sense for failure demand in a health setting. To do this, numerous methods were considered, some were trialled, and others were rejected outright. This section is a part of the justification for the methods that were ultimately used. In many instances, convenience factors were definitive in the ultimate choice of the methods used in this thesis.

"Failure to do something or do something for the customer" [44] is not failure demand. This moment at which a customer leaves a service system unsatisfied is only the beginning. It is only when this customer returns to the system, to consume the same resources of the system as before – in pursuit of the incomplete resolution – that it becomes failure demand. This presents considerable difficulty for measuring failure demand, because without a "memory" of the previous interaction, and the unsuccessful resolution, failure demand cannot be identified. Ensuring that this memory exists was one of the more difficult tasks for this study.

Several models were considered, some rejected, and others attempted and later abandoned. These approaches all aimed to establish a value for failure demand. This section highlights the methods that are not in use in this thesis, they include patient surveys or interviews, data mining admissions data from a large hospital, an observational study in an operating theatre and records analysis in radiology.

3 - 3.1. Approach 1: Finding the memory in patients.

Description

The study was to be conducted in the queue of an outpatient department (OPD) of a public hospital. Patients would be interviewed in person to establish whether they are there because of previous system failure and could therefore now be classified as failure demand.

Advantages

Rich and deep qualitative data could be extracted from the interviews with patients, delivering many additional data points and more information on the lived experience of people experiencing failure demand.

Disadvantages

The disadvantages to this approach were overwhelming.

Firstly, The Human Research Ethics Committee of the home institution has been highly resistant to permitting non-clinical researchers (such as industrial engineers) to engage directly with patients. In South African public health care – patients are generally classed as "vulnerable" [262] and gaining true informed consent would have been a matter of considerable difficulty [262].

Secondly, the nature of failure demand is complex [33], if managers in the English NHS were unable to fully understand the complexities of the concept [194], it is uncertain whether lay-people could adequately grasp the concept in a short amount of time to ensure that the quality of data recorded from the participants was accurate. Differentiating between "having been at the OPD before" and "having been at the OPD before *for the same unresolved reason*", was challenging. This was even further exacerbated by a common observation in the experience of the author that patients in South African public hospitals are often not informed what their condition is, but merely instructed to take medication. In instances where patients would know their medical history, this would yet again have been a barrier to access as the ethics board would have had concerns with engineers taking patient histories [262].

Thirdly, this method would have been resource and time intensive. To do justice to collecting the data, one-on-one interviews would have been necessary. This should have been in a separate room and could not have been conducted in the general waiting area of the hospital and may have required the services of translators. This approach would further have required considerable time to ensure justice is done to explanations and data collection and processing would only have been for a small subset of patients at a particular hospital. The return on effort related to data collected was misaligned with the objectives for this study.

Reason for rejection

Access to patients would have been limited by the ethics board, patients in the system would not serve as adequate "memories" for previous system engagements. And the resources required to do justice to the data collection requirements led to the decision to abandon this approach.

3 - 3.2. Approach 2: Data mining admission records of a public hospital

Description

The team was given access to the admissions data set from a large public hospital. An analytical tool was developed to mine this data set to assess failure demand. The logic employed was that the unique patient ID would be tracked through the system and all repeat interactions would be cross referenced against diagnostic codes (ICD10). A rules-engine would be used to assess whether an interaction can reasonably be interpreted as failure demand.

The data set contained approximately 902 000 patient entry records. For this study, a "robot" was programmed that "accompanied" each patient through their multiple entrances into the hospital over a period of eighteen months. It was intended to see the instances where the same person reappeared presenting with the same ICD10 implying previous non-resolution.

This method was similar to work done in call centres [44, 110, 185]. In these instances, a repeated reference number was interpreted as failure demand. The question whether a returning patient, with the same diagnostic code was in every instant failure demand required deeper consideration.

Advantages

The use of ICD10 codes would allow non-failure demand repeat interactions to be excluded. For example, somebody presenting with one condition, and a completely different one at a later stage would not be deemed failure demand. Similarly, based on a rules-engine (which developed into the framework presented in Chapter 4) chronic care, for example renal care, would not be interpreted as failure demand. This required a recognition that resolution in health care is not as simple as in other services. Therefore the work by Lillrank [241] differentiating between demand for care and demand for cure.

The analytic tool has been developed and is available to analyse similar databases at other hospitals or at other times to assess failure demand levels and to identify opportunities for improvement. This tool can be used to create snapshots of failure demand which is the appropriate way to use failure demand [44].

The scale of the data means that good, long-term and diachronic analyses are possible. This data set could deliver a comprehensive insight into failure demand and could be modified to show failure demand levels over time, and if matched to a suitable improvement strategy, could also be used to show opportunities for system improvement.

Disadvantages

The data received for this study were of very low quality. The data were not full medical records, but merely hospital entry/admission records. The ICD10 codes were often omitted (68 % of cases did not have diagnostic information) and those that were captured seem to have been captured by non-clinical clerks, frequently only capturing “emergency treatment”. Although the records always captured a patient ID and an entry date, discharge dates were present in only the minority of cases. This also meant that diagnostic information was not updated by clinically trained people.

Reasons for rejection

Unfortunately, this study's results had to be discarded due to the poor quality of data and the resultant, almost meaningless, conclusions based on the inadequacy of the data collected.

The tool has been completed and is mature and stable. As such, it remains useful, assuming that better, more complete data sets can be found. It can be applied to assess failure demand for other settings.

3 - 3.3. Approach 3: Making use of cloud-linked devices.

Description

Outpatient departments would be provided with Android-based tablet computers. Upon patient departure prior to resolution, nursing staff would record the exit on a custom, single-touch, graphical user interface, noting unresolved exit, and a menu of standard reasons for the exit. Data collection, storage and processing could be automated. One prototype was developed and worked.

Advantages

This system would record reasons for system exit, whilst in most other abandoned methods, the reasons would be inferred from context.

The custom app could provide rich and live data and if use were consistent, a rich picture of system failures could emerge.

Disadvantages

This system would not measure failure demand, only system exits before resolution. A further concern is that it was unlikely that the staff who will be asked to use the system would do so consistently, with a particularly problematic gap over shift changes and forgetfulness during particularly busy periods, which would probably correspond to the periods of maximal system failure. The tablet devices also represented a risk due to loss, theft, or damage. The tablets were expensive and development of the software to a more user-friendly product would have been time consuming.

Reason for rejection

The high cost, risk of non-use and the fact that this system would not measure failure demand (but rather system failure) led to the rejection of this method.

This remains a study worth executing, as failure demand should always be referred to systemic root causes, and this approach would have identified these more capably than other approaches.

Conclusion

Several methods were considered for understanding different elements of failure demand. Although each had their strengths, the final decision to base the empirical element for this thesis on a pharmaceutical supply-chain derived from the fact that the data available in this case was of high quality and complete, and that the findings from that would be more trustworthy and that therefore the conclusions would be more valid.

3 - 4. Structure of the study

This study is structured to assess the prevalence of failure demand. To do so, the concept of system failure will be introduced, and the root causes examined. This approach is appropriate for the exploratory nature of the study, and addresses the theory building earlier phases and the methodological expansions.

To do this, a framework describing demand modalities is developed from grounded sources, observations and experience (Chapter 4). Out of this a functional mental model emerges to identify the behaviours that describe failure demand in health care (Chapter 5). One of the drivers for failure demand, "Inventory and supply-chain management", is then tested by data mining the order and delivery records for the national pharmaceutical supply-chain of a developing country (Chapter 6).

This chapter reveals the high-level structure and methodological nature of the study. More detailed methodological issues are described in the individual papers, as necessary.

Figure 3-12 (reproduced in larger size in Appendix 2 – Page 336) summarises the methods employed in this thesis. It shows the high-level methodological structure, how these elements are broken up into discreet deliverables and the sources of data and their processing requirements. As the work for this thesis made use of several approaches, this section shows the overall study structure allowing for the integration of outputs to address the research questions.

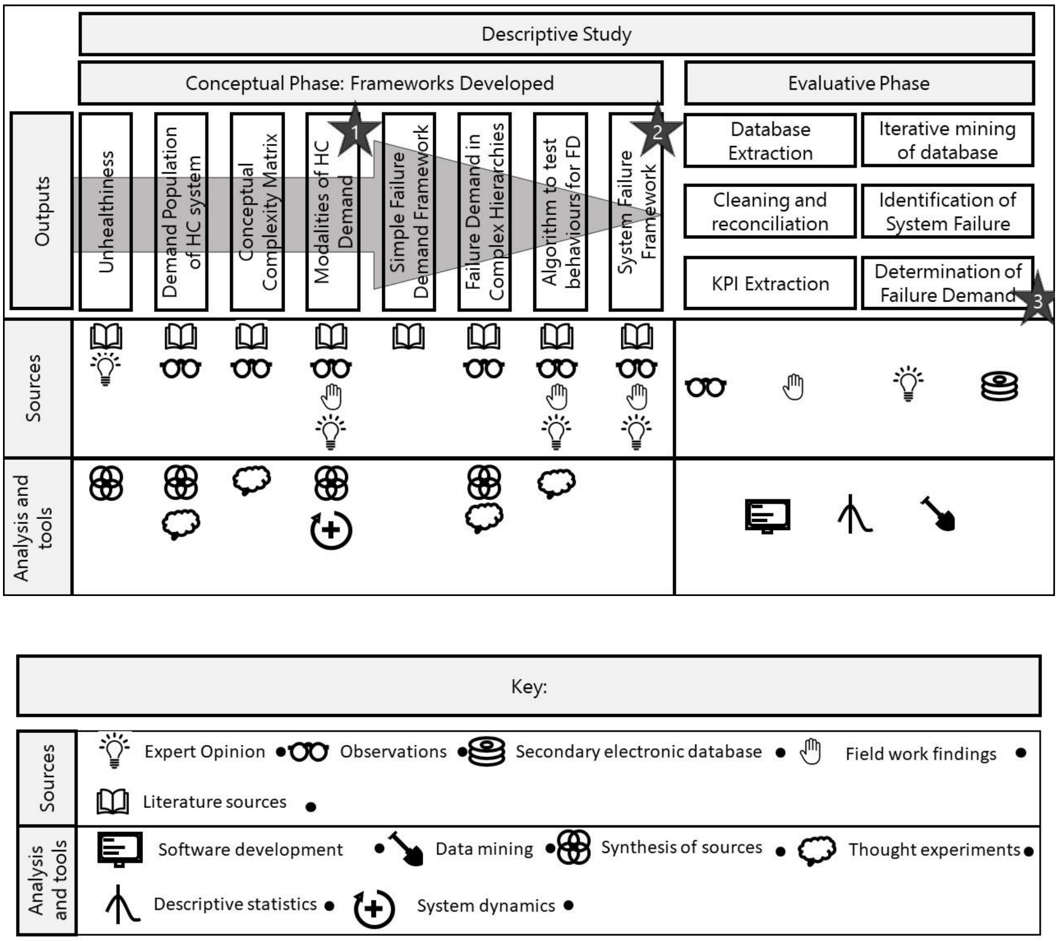


Figure 3-12: Overall study structure

This descriptive study is structured into two key phases. Firstly, a phase during which two important apex – conceptual-frameworks are developed. These are highlighted with stars in Figure 3-12 and address the first two research objectives in Section 1 - 2.5 .

In the second, Evaluative - phase, these frameworks are used as the basis for assessing the central research question, deriving a baseline value for the incidence of failure demand in a selected health care setting. The phases lead into each other, and there is an overall dependence between the phases to establish a coherent narrative.

3 - 4.1. Milestones

The key output milestones for this thesis are shown in numbered stars in Figure 3-12. These milestones represent the papers that were written for this thesis and address the objectives for this study. Working backwards[10], the key questions that required answering[11] were:

③ What impact does failure demand have on service delivery (in a selected health care setting)? To answer this, it was necessary to have a sense for ② what failure demand and system failure look like in a health care setting. To understand failure demand, it is important to ① clarify the demand ecosystem in this sector.

The data sources and analytical approaches for the two phases will be discussed separately in Sections 3 - 5 and 0 as their natures are distinct. The Conceptual phase creates theoretical frameworks that enable the eventual measuring of *failure demand* in health care, whilst the Evaluative phase is a data-driven test of the phenomenon that is described in phase one.

Section 3 - 5 will discuss the Conceptual phase, the data sources and analytical techniques used for these, whilst Section 3 - 6 will deal with the same features of the Evaluative phase.

[10] Which is why the numbering below appears to be in reverse order.
[11] Reworded from 1 - 2.4 and 1 - 2.5 on page 32.

3 - 5. Conceptual Phase

The conceptual parts of this study allow milestone one and milestone two to be reached (Figure 3-12 and reproduced larger as Appendix 3 on Page 337). This part of the research made use primarily of literature sources, and field work, over a decade of exposure to the health care industry and included direct observations, project work and expert opinion gathered during this time.

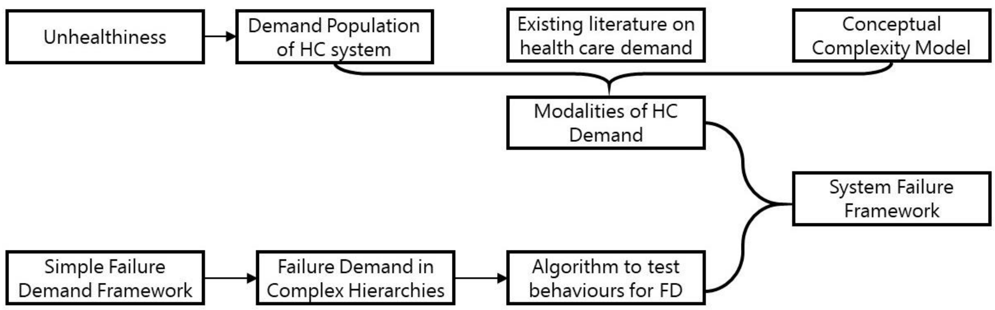


Figure 3-13: Conceptual phase structure

Figure 3-13 represents the high-level structure for the Conceptual phase and shows how the elements built up towards creating a model that could be used to identify system failure in a health care setting.

The first milestone was the creation of a model of health care demand modalities and required an exploration of the natural demand population in a health care setting. This knowledge was augmented by current thinking on demand structure, from literature. A model that deals with complexity as a demand modifier was added, to create an overall idea of how demand is shaped in health care. This model is captured in the first paper of this thesis.

The second paper narrows the focus to failure demand, exploring first failure demand in familiar contexts, and then enriches it by adding complex hierarchies as a modifier. This paper contextualises the subtleties that exist between the closely related ideas of system failure, and failure demand.

This creates a framework that can be used to establish whether a phenomenon is system failure or not. The model obtained is the baseline for further study and is used as the basis for the Evaluative phase described in Section 3 - 6.

3 - 5.1. Data sources.

3 - 5.1.1. **Literature sources**

Literature sources used were consulted in line with the principles described in Section 2 - 1.1 – the selection of literature sources.

3 - 5.2. Field work

3 - 5.2.1. **Pilot studies and field work**

The pilot studies that framed a departure point for the work in this thesis were described in section 1 - 1.1. with some methodological elements strengthened in Section 3 - 3. This created context and provided benchmark data, which was used and referred to in the remainder of this thesis.

This thesis contains direct observations [16] that were made during the course of the work. Beyond this, the thesis also makes use of the researcher's observations spanning more than a decade of work in a health setting. This period of experience includes nine peer reviewed published articles in this field. Additional observations were made in numerous un-published undergraduate- and Masters-projects that were supervised by the researcher in the health sector.

The remaining data sources from Figure 3-12 are predominantly in the Evaluative phase and are discussed in Section 3 - 6.1.

3 - 5.3. Analysis approaches

3 - 5.3.1. **Thought experiments.**

Practitioners of the natural sciences are frequently dismissive of thought experiments as a means for scientific discovery [263]. Thought experiments are not merely what we think – of which Neil De Grasse Tyson is highly critical[264]. Critics especially lament that the wording seems to add the veneer of legitimacy to "thought", by including the word "experiment" to the drama in the human mind [263].

By contrast, Norton believes that thought experiments are essentially equivalent to argument [263] – which are both important mental constructs. He believes that the legitimacy of thought experiments can be conflated with an argument which is central to the scientific method of discovery. Perhaps the greatest true difference then is the picturesque nature of thought experiments.

However, as Sorensen points out [265], some of history's important scientific contributions come in the form of thought experiments which include Schrödinger's cat, Einstein's Elevator (or train)[12], Newton's bucket and Maxwell's Demon [265]. Sorensen goes on to clarify the nature of a thought experiment as the act of placing a concept "under the stress of bizarre situations" [267 Pg. 135] whilst Brown calls Thought Experiments "laboratories" albeit "of the human mind" – suitable for the natural sciences [268 Pg. 1].

As a moderating view, Di Paolo et al. describe simulation, a trusted part of modern science, as "opaque thought experiments". Taking the experiment outside of the human mind, even if the essential thought processes, according to which the simulation runs, is constructed no differently than had the simulation taken place in an armchair [269 Pg. 497].

The nature of thought experiments becomes a philosophical one – it can be a reformulated argument, as Norton [263] claims, and therefore intrinsically scientific, a simulation [269] or it can be a laboratory, the view that Brown and Sorensen favour [265, 267, 268]. Regardless of their positions, these authors support the scientific usefulness of thought experiment and its place in discovery.

This thesis makes use of thought experiments. It asks questions such as "what happens if a healthy person goes to a hospital?", "what happens if a sick person does not?". "What is the effect, If I arrive at the hospital for care, and the surgeon is having lunch", or "what happens, if a well running

[12] It is interesting to read Norton's paper [266] that treats several of Einstein's thought experiments and follow his reasoning that tries to show that Einstein's thought experiments tend to be reformulated, picturesque arguments.

hospital suddenly has to deal with a bus accident. How aggrieved can I as a patient with a broken arm be that I am delayed?".

And although thought experiments are used as picturesque ways of sharing ideas – this thesis ensures that the entirety of the discovery is not conducted in the thought laboratory only.

3 - 5.3.2. **Synthesis**

As Di Paolo says, thought experiments are inspired by the context of the researcher [269]. An entomologist might use pictures from a termite colony to construct the image [269]. For this thesis, the context from which images were drawn, was from the researcher's experience in hospitals and often uses individualised vignettes, and though these may seem like personal stories, they are the synthesised products of discovery.

Although thought experiments (or argumentation) was an important part of the construction of the first two papers this was inadequate without a grounded basis from literature, which is an important constituent of Norton's [263] argumentation and that of Di Paolo [269] – that the thought experiment works when it is a component of a synthesised whole – that includes text sources, observations, thought experiment and argumentation and the ultimate synthesis of many sources.

The synthesis of sources is mostly shown in the papers, with the sequence of source, inspiration, observation and mechanism highlighted.

3 - 5.3.3. **System dynamics and causal loop diagrams**

Causal loop diagrams were chosen to represent the relationships between demand-shaping elements (see Section 4 - 5.3 on page 169). This primer introduces the reader to causal loop diagrams and a suitable approach for constructing them.

The Conceptual phase of this research primarily sets out to understand the relationships between various factors. This is typical for the construction of a conceptual framework [270], or as Jabareen says, "network, or a plane," of interlinked concepts that together provide a comprehensive understanding of a phenomenon or phenomena" [271 Pg.50]. But whilst Jabareen's conceptual frameworks do not explore causality, the discipline of system dynamics attempts to bring this dimension to the relationships in theoretical models [99].

Causal loop diagrams are a way to represent these relationships pictorially but also analytically. The key idea underlying this – is an understanding of the causation vector in a relationship [99].

Causal loop diagrams represent the closed-loop relationship between two phenomena. They require causation – in other words, one phenomenon causes change in the other [272]. The type of causation can be a positive relationship – i.e. a positive change in the driver results in a positive change in the driven, or it can be a negative relationship – in which a positive change in the driver results in a negative change in the driven [99]

Two prototypical loops exist – the reinforcing loop, in which a phenomenon is continuously strengthened (or weakened) and a balancing loop, in which a phenomenon plateaus [99, 273].

Figure 3-14 was developed for this thesis and shows two familiar introductory phenomena to illustrate the idea of Causal Loop Diagrams; reinforcing loops and balancing loops.

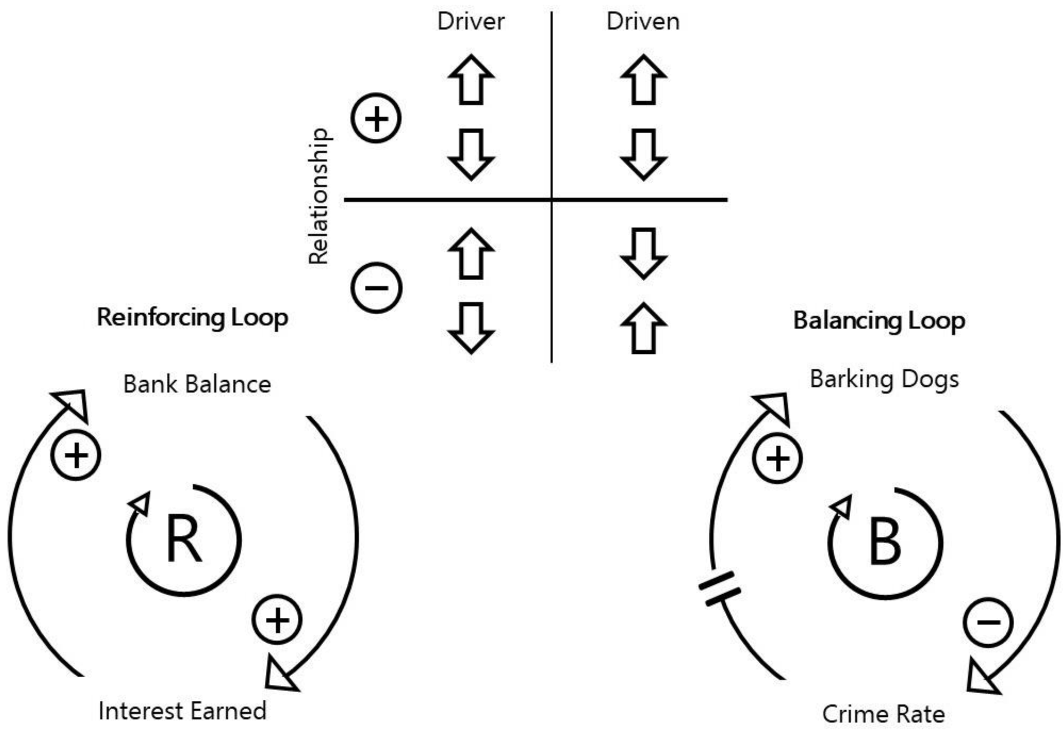


Figure 3-14: Causal Loop Diagrams Explained [99, 274]

Reinforcing loop

In the reinforcing loop example (which has been used extensively, for example by Brzezina [274], bank balance is positively associated with interest earned. The more money is in the bank, the greater the amount of interest earned (or conversely, the greater the deficit, the higher the interest on debt, owed to the bank).

The interest earned raises the bank balance (again positively associated), leading in turn to higher interest for the next cycle. This escalation (in the absence of external phenomena) leads to a compounding of interest, and a continuously growing bank balance (or deficit).

Balancing Loop

By contrast, the thought experiment presented in the balancing loop example, shows how the relationship between dogs and crime works. The idea is that the number of barking dogs, frightens off criminals (indicated by use of the negative association). Thus, the more barking dogs there are, the less crime will exist. However, as crime rates go down – dog ownership might follow that trend (indicated with the positive association). As a result, there will be fewer barking dogs (in time – the time delay shown by the bisected relationship line) and opportunist criminals will enter the profession – yet again, conforming to the negative association. Of course, as a result, citizens will again "invest" in dogs, thus leading to an eventual decline in crime. As there is an implied stasis in this system, it is considered a balancing loop, indicated by the "B" at the loop's centre.

Approach

To construct causal loop diagrams for this thesis, Haraldsson's six key steps were followed [275]. (1) First the phenomena were identified, (2) after this, the relationship was assessed between the phenomena, (3) and the feedback path for secondary causality was assessed. After this (4), the polarity for the first link was determined, the polarity for the feedback assessed (5), and finally, based on the closed loop with completed polarity, the type of loop was identified (6). This approach was expanded to add secondary and concentric loops.

3 - 6. Evaluative Phase

The Evaluative phase was conducted in part as a field work deployment with an associated secondary data source analysis. During this phase, the ideas emerging from the Conceptual phase were tested against a real-world scenario. The Evaluative phase was run as a mixed methods study [276]. Observations were made in the country, which formed the contextual basis for the analytical findings from the empirical study. The analytical findings emerged from a data mining, processing and modelling exercise, making use of secondary data sources provided by the ministry of health in the country under review.

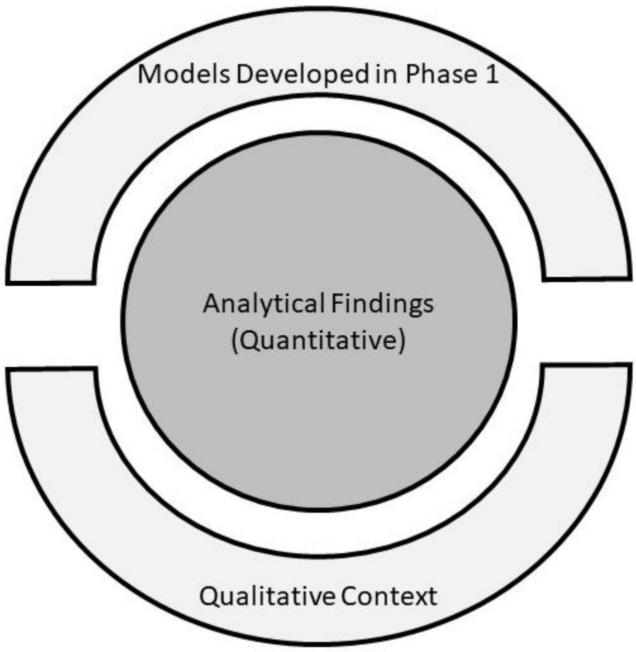


Figure 3-15: Structure of the Evaluative phase

Figure 3-15 represents a mental model of how this phase is structured, making use of the observations as a qualitative context to support the analytical findings, and making use of the conceptual models, developed in Phase one as the theoretical underpinnings for the analysis.

Although the findings of the paper are presented largely based on the analytical results, the qualitative assessments allowed for a contextual understanding, that enriched the case study [16], and allowed for the making of realistic and useful conclusions.

3 - 6.1. Data sources for the evaluative phase

3 - 6.1.1. Direct and participant observations

This part of the study was conducted in the national pharmaceutical supply-chain in a developing country. Four visits were made to the country and to the supply system over the course of six months, with a total of eighteen days spent doing direct observations, following the philosophy of "Going to the Gemba" [277], and taking inspirations from Roser [278], Goodson[13] [279] and Hopp & Spearman[14] [280]. During the field work portion, extensive notes were made and kept of observations and discussions with global experts in the fields of pharmaceutical supply-chains.

Ethnographic research makes use of a technique known as Participant Observation [17], in which the researcher gets deeply immersed in the organism that they are investigating. In many cases living with and, as one of the members of, the society being studied [281].

[13] Read a plant fast
[14] Factory physics

The nature of the field work should rather be interpreted as Direct Observation [282]. This is possibly one of the oldest research techniques in science, and mimics the scientist in a laboratory, observing, but not altering outcomes [282]. Which is again aligned with an overall positivist philosophical position. Given the limited duration of the field work, and the depth of immersion in the ecosystem of the workplace, it is fair, to rather interpret the means of observation as direct observation.

Although the depth of context achieved by these methods might be reduced when compared to participant methods, the bias inherent to participant methods is reduced with greater distance and researcher objectivity [283].

3 - 6.1.2. **Expert opinion**

During the time spent researching the health system, the researcher was in direct contact with five international supply-chain consultants, each of whom had a career of at least twenty years of exposure to global supply-chain best practices. The unstructured interviews with these authorities led to contextualising findings, particularly creating context for verifying best practice benchmarks and understanding the limitations of the extant systems. Extensive notes were recorded in a notebook, which was carried during the entire field work phase.

3 - 6.1.3. **Secondary database**

The quantitative assessment was conducted on databases that were produced by the agency that manages pharmaceutical distribution in the country being investigated. These databases contain vast data that is mostly automatically captured logged and stored. With the support of the national ministry of health, this database was made available for this research. The database was sliced to provide only necessary information as the databases contain significantly more fields than are relevant for this study. Permission to make use of this database was provided by the director general of the relevant agency.

“Big data” is identified as a “buzzword”. “Big Data” is ill-defined and unclearly delimited as a concept [284, 285]. Nevertheless, some convergence exists in the academic literature that Big Data must be “Too Much, Too Fast and Too Hard” [284 Pg. 4] or as Madden more soberly frames it as the three Vs – Volume, Velocity and Variety [284], and all of this on the petabyte scale [285].

By this measure, the data that were used in this study are not “Big Data” but nevertheless substantial. The data represented several hundred thousand transactions at a time, that spanned up to three years. As approximately twenty further datapoints were recorded per transaction, the dataset contained several million points, that had to be cross referenced and cross mapped across multiple relational and hierarchical databases.

3 - 6.2. Analysis approaches

The key method of analysis for this phase, was quantitative, making use of software, custom developed by the author, for this study. This software was used to analyse the system by first cleaning and then processing and analysing the data set. From this, important static indicators were extracted from the set, which was followed by a specially developed dynamic routine that was able to identify system failure, allowing for conclusions about failure demand to be drawn.

3 - 6.2.1. **Creation of custom analytical software**

The methods used to process the data could be thought of as data mining approaches from first principles, the approaches could also be seen as modelling as well as data science.

The software was developed for this study, by the author using Visual Basic for Applications (VBA), a programming language that integrates with Microsoft Excel. The advantage of this approach is that the database as well as the code are accessible to all who make use of MS Excel – which includes the research site.

3 - 6.2.1.1. Platform

Microsoft Excel was used for this project for three key reasons. Firstly, Excel is ubiquitous – meaning that it is loaded on most computers [286], which includes those of the agency being investigated. This means that in future similar analyses can be run by the in-country team in a familiar ecosystem.

Secondly, Excel through its integration with VBA is immensely powerful at data processing, surpassing the native abilities of Excel alone significantly, matching the native processing abilities of any object-oriented programming language joined to Excel's capabilities as a relational database. In fact, making use of VBA is generally able to overcome the rigidness of Excel analysis entirely. This integration also provides the user with native Excel visualisation and worksheet tools that may not be as accessible in other platforms [286].

Thirdly, to make use of Microsoft Excel for this analysis is a proof of concept to show that these kinds of analyses can be done, making use of standard software available on most computers. The advantage of this is that researchers with limited resources interested in similar types of analyses may be confident to make use of this platform, rather than pausing or abandoning an analysis because the acquisition cost of specialised data mining software can be prohibitive. This means that access to sophisticated software can be less of a constraint.

3 - 6.2.1.2. Approach

The overall approach followed when writing the code, was to automate the entire analysis process. As a result, analyses are repeatable, and new data sets can be analysed at once, without the need for cumbersome (and prone to error) human processing.

The approach was to create a highly modular program that allowed for individual processing steps to be added or removed through references in the main code.

To achieve this – all the steps required for the analysis were mapped, and code was developed to carry out such processes. This included all the pre-processing the processing and the results presentation and reporting [287-290].

During the pre-processing, the data were cleaned – duplicates were removed, consistent naming conventions were applied, keystroke and other errors were removed or corrected, and data were cross checked for underlying consistency.

In the analysis phase, the traditional KPIs were extracted, and the iterative identification of system failure events undertaken. This was further mapped against instances of returned business – which corresponds to instances of failure demand.

The final phase automated the presentation of results and suitable visualisations.

Initial work made use of the spreadsheets themselves to store values, however, working with large datasets this meant a complete analysis run lasted approximately 65 minutes. This proved cumbersome, so the code was redeveloped making use of arrays stored in computer memory. After the redevelopment, the analysis completed in approximately six and a half minutes – roughly one tenth of previous computation time.

3 - 7. Verification and validation

Establishing the validity of the theoretical models is achieved mainly by ensuring the internal validity of the artefact [20]. This is done by ensuring that the approach was suitable, grounded and well-reasoned, and that the data source selection is appropriate and sufficiently broad.

In the case of the frameworks, the data sources are varied, as shown in sections 3 - 5.1 and 3 - 6.1 and thus presented the researcher with a diverse set of perspectives from which to build the models. This included grounded literature sources [291], modified by field work, observations, experience and other sources. This triangulation of inputs led to synthesised products that are plausible.

Construct validity [20] assesses whether the logical build-up of artefacts and results delivers expected results. This can be evaluated by tracing through a sequence of justified arguments. This can make iterative use of expert opinion leading to a validated documented construction.

These models are artefacts that can be used to understand the demand patterns for health systems. Richer external validity will be gained in time as more researchers and practitioners make use of the frameworks to evaluate system performance. It is anticipated that these artefacts will be modified in time. The evaluative phase of this study tests the underlying logic in these models. The results from the Evaluative phase validate the models; because the final output of a failure demand value is range validated – comparing realistically to existing failure demand benchmarks from literature, whilst the simple KPIs map closely to those computed by the existing WMS/ERP – system.

Quantitative data were verified by exploring the method of capture. The automated data collection by the ERP system allowed for minimal error – and the major concerns, keystroke errors and inconsistent naming conventions could be automatically corrected, leaving a unified database. The use of a consistent repeatable tool supported the internal validity of the findings due to the standardisation of the analysis [20]

The external validity was supported using the full available dataset. Thus, avoiding selection biases or the inadvertent exclusion of outlier data.

Unobtrusive data collection methods [292] were used to limit the impact of the researcher on the external validity [293] of the work by reducing novelty effects [20]. It is agreed that the researcher can play a role in the modification of results, leading to reduced validity. This can be managed through triangulation of source, [20] or by reducing the power that the researcher has over research findings.

Making use of direct observations (rather than the more immersive participant methods) and secondary data sources, the prominence of the researchers was reduced, meaning that behaviours were likely not materially modified, thus limiting the impact of observers' paradox [294]. Although some disagreement exists whether the Hawthorne Effect is real – it remains a great illustrative tool to show the inadvertent impact of observation on human behaviour and system performance and efficiency [295]. Due to this, the external validity of the work is strengthened, leading to greater generalisability [20].

3 - 8. Ethical Considerations

The field work elements of this thesis were conducted after applying for ethics clearance from the Human Research Ethics Committee (Medical) at the University of the Witwatersrand, Johannesburg.

The two pilot studies cleared for research with reference numbers:

M160565 and M160566

And the main study was conducted under

M170744

Limitations placed on research by the ethics clearance

No onerous limitations were placed on this research by the wording of the ethics clearance received, However, this was in part because the author had previously had difficulty in receiving ethics clearance in cases where access to patients was sought. As a general principle, the Human Research Ethics Committee is resistant to granting access to patients to non-clinically trained researchers, because so many cohorts are deemed vulnerable in South Africa [262]. For that reason, this study was purposefully scoped to exclude the need to interact directly with patients, but rather requested access to data recorded by hospital admissions systems and ERP systems for pharmaceutical supply-chains.

3 - 9. References

16. Yin, R.K., *Case Study Research: Design and Methods*. 1987, Sage Publications, Inc., Beverley Hills
17. Levine, H.G., Gallimore, R., Weisner, T.S., and Turner, J.L., *Teaching Participant-Observation Research Methods: A Skills-Building Approach.* Anthropology & Education Quarterly, 1980. 11(1), p. 38-54.
20. Marczyk, G., Dematteo, D., and Festinger, D., *Essentials of Research Design and Methodology*. 2005, John Wiley & Sons Inc, Hoboken, New Jersey
31. Crotty, M., *The Foundations of Social Research: Meaning and Perspective in the Research Process*. 1998, Sage, New York City
33. Seddon, J., *Failure Demand.* 2020.
44. Seddon, J., *Freedom from Command and Control: Rethinking Management for Lean Service*. 2005, CRC press, Boca Raton, Florida
99. Senge, P.M., *The Fifth Discipline.* Measuring Business Excellence, 1997. 1(3), p. 46-51.
110. Piercy, N. and Rich, N., *Lean Transformation in the Pure Service Environment: The Case of the Call Service Centre.* International Journal of Operations & Production Management, 2009. 29(1), p. 54-76.
185. Marr, B. and Neely, A., *Managing and Measuring for Value: The Case of Call Centre Performance*. 2004, Cranfield School of Management and Fujitsu, Cranfield, UK.
194. Walley, P., Found, P., and Williams, S., *Failure Demand: A Concept Evaluation in Uk Primary Care.* International Journal of Health Care Quality Assurance, 2019. 32(1), p. 21-33.
241. Lillrank, P., Groop, P.J., and Malmström, T.J., *Demand and Supply–Based Operating Modes—a Framework for Analyzing Health Care Service Production.* Milbank Quarterly, 2010. 88(4), p. 595-615.
248. Armstrong, D.M., *Belief, Truth and Knowledge*. 1973, CUP Archive
249. Seth, A., *Epistemology and Ontology.* The Philosophical Review, 1894. 3(5), p. 568-582.
250. Armstrong, D.M. *A Naturalist Program: Epistemology and Ontology.* in *Proceedings and Addresses of the American Philosophical Association*. 1999. Newark, Delaware, JSTOR.

251. Raddon, A., *Early Stage Research Training: Epistemology & Ontology in Social Science Research.* Generic Skills Training for Research Students, 2010, p. 1-14.
252. May, T. and Williams, M., *Knowing the Social World*. 1998, Open University Press, Maidenhead, UK
253. Berryman, D.R., *Ontology, Epistemology, Methodology, and Methods: Information for Librarian Researchers.* Medical Reference Services Quarterly, 2019. 38(3), p. 271-279.
254. Al-Saadi, H., *Demystifying Ontology and Epistemology in Research Methods.* Research Gate, 2014. 1(1), p. 1-10.
255. Saunders, M., Lewis, P., and Thornhill, A., *Research Methods for Business Students*. 8th ed. 2019, Pearson Education Ltd, Harlow, United Kingdom
256. Moon, K. and Blackman, D., *A Guide to Ontology, Epistemology, and Philosophical Perspectives for Interdisciplinary Researchers*. 2017.
257. Goertz, G. and Mahoney, J., *Concepts and Measurement: Ontology and Epistemology.* Social Science Information, 2012. 51(2), p. 205-216.
258. Cresswell, J.W., *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. 2009, Sage, Thousand Oaks
259. Lawson, T., *Comparing Conceptions of Social Ontology: Emergent Social Entities and/or Institutional Facts?* Journal for the Theory of Social Behaviour, 2016. 46(4), p. 359-399.
260. Moon, K. and Blackman, D., *A Guide to Understanding Social Science Research for Natural Scientists.* Conservation Biology, 2014. 28(5), p. 1167-1177.
261. Ahmed, A., *Ontological, Epistemological and Methodological Assumptions: Qualitative Versus Quantitative.* Research Gate, 2008.
262. Moodley, K. and Myer, L., *Health Research Ethics Committees in South Africa 12 Years into Democracy.* BMC Medical Ethics, 2007. 8(1), p. 1.
263. Norton, J.D., *Are Thought Experiments Just What You Thought?* Canadian Journal of Philosophy, 1996. 26(3), p. 333-366.
264. De Grasse Tyson, N., *Scientific Thinking*, in *Neil De Grasse Tyson Teaches Scientific Thinking and Communication Skills*. 2019, Yanka Industries,, San Francisco USA.
265. Sorensen, R.A., *Thought Experiments*. 1998, Oxford University Press on Demand, Oxford, UK

266. Norton, J., *Thought Experiments in Einstein's Work.* Horowitz and Massey, 1991. 1991, p. 129-148.
267. Brown, J.R., *Review of Thought Experiments by Roy Sorensen.* Canadian Journal of Philosophy, 1995. 25(1), p. 135-142.
268. Brown, J.R., *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*. 2011, Routledge, New York City
269. Di Paolo, E.A., Noble, J., and Bullock, S., *Simulation Models as Opaque Thought Experiments.* Univerisity of Southampton, 2000.
270. Forrester, J.W., *Counterintuitive Behavior of Social Systems.* Theory and decision, 1971. 2(2), p. 109-140.
271. Jabareen, Y., *Building a Conceptual Framework: Philosophy, Definitions, and Procedure.* International Journal of qualitative Methods, 2009. 8(4), p. 49-62.
272. Sterman, J.D., *Business Dynamics: Systems Thinking and Modeling for a Complex World*. 2000, McGraw Hill, Homewood, Illinois
273. Sterman, J., *Business Dynamics*. 2010, Irwin/McGraw-Hill, Homewood, Illinois
274. Brzezina, N., Kopainsky, B., and Mathijs, E., *Can Organic Farming Reduce Vulnerabilities and Enhance the Resilience of the European Food System? A Critical Assessment Using System Dynamics Structural Thinking Tools.* Sustainability, 2016. 8(10), p. 971.
275. Haraldsson, H.V., *Introduction to Systems and Causal Loop Diagrams.* System Dynamic Course, Lumes, Lund University, Sweden, 2000.
276. Bulsara, C., *Using a Mixed Methods Approach to Enhance and Validate Your Research.* Brightwater Group Research Centre, 2015, p. 1-82.
277. Womack, J., *Gemba Walks: Expanded 2nd Edition*. 2013, Lean Enterprise Institute, Cambridge, Massachusetts
278. Roser, C., Lorentzen, K., and Deuse, J., *Reliable Shop Floor Bottleneck Detection for Flow Lines through Process and Inventory Observations: The Bottleneck Walk.* Logistics Research, 2015. 8(1), p. 7.
279. Goodson, R.E., *Read a Plant-Fast.* Harvard Business Review, 2002. 80(5), p. 105-113.
280. Hopp, W.J., *Single Server Queueing Models*, in *Building Intuition*. 2008, Springer. p. 51-79.

281. Jorgensen, D.L., *Participant Observation.* Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource, 2015, p. 1-15.
282. Jersild, A.T. and Meigs, M.F., *Chapter V: Direct Observation as a Research Method.* Review of Educational Research, 1939. 9(5), p. 472-482.
283. Iacono, J., Brown, A., and Holtham, C., *Research Methods--a Case Example of Participant Observation.* Electronic Journal of Business Research Methods, 2009. 7(1).
284. Madden, S., *From Databases to Big Data.* IEEE Internet Computing, 2012. 16(3), p. 4-6.
285. De Mauro, A., Greco, M., and Grimaldi, M. *What Is Big Data? A Consensual Definition and a Review of Key Research Topics*. in *AIP conference proceedings*. 2015. American Institute of Physics.
286. Sengupta, C., *Financial Modeling Using Excel and Vba*. Vol. 152. 2004, John Wiley & Sons, Hoboken, New Jersey
287. Guttag, J., *Introduction to Computation and Programming Using Python: With Application to Understanding Data*. 2016, MIT Press
288. Bird, J., *Science for Engineering*. 2015, Routledge
289. Zelle, J.M., *Python Programming: An Introduction to Computer Science*. 2004, Franklin, Beedle & Associates, Inc.
290. Shaw, Z.A., *Learn Python 3 the Hard Way: A Very Simple Introduction to the Terrifyingly Beautiful World of Computers and Code*. 2017, Addison-Wesley Professional
291. Barth, A., Caillaud, E., and Rose, B. *How to Validate Research in Engineering Design?* in *DS 68-2: Proceedings of the 18th International Conference on Engineering Design (ICED 11), Impacting Society through Engineering Design, Vol. 2: Design Theory and Research Methodology, Lyngby/Copenhagen, Denmark, 15.-19.08. 2011*. 2011.
292. Rosenthal, R. and Rosnow, R.L., *Essentials of Behavioral Research: Methods and Data Analysis*. Vol. 2. 1991, McGraw-Hill New York
293. Bracht, G.H. and Glass, G.V., *The External Validity of Experiments.* American Educational Research Journal, 1968. 5(4), p. 437-474.
294. Shanmuganathan, T., *Ethics and the Observer's Paradox.* relc Journal, 2005. 36(1), p. 73-84.
295. Wickström, G. and Bendix, T., *The" Hawthorne Effect"—What Did the Original Hawthorne Studies Actually Show?* Scandinavian Journal of Work, Environment & Health, 2000, p. 363-367.

Chapter Four: **Demand Modalities in Health Care**

4 - 1. Abstract

Objective

This study addresses the gap in defining, understanding and measuring demand in health care, and its drivers – with potential improvements in health productivity.

Method

An exploratory-conceptual study synthesising literature sources, expert-experience, and -observations in health system improvement.

Results

We make three contributions. Firstly, a mental model defining the demand – building blocks. Secondly, we introduce two new demand classes: "escalation" (demand due to delayed treatment) and "false" demand (originating in the healthy population). And thirdly a complex demand measurement framework. Results are presented as an equation, a conceptual model and a system dynamic model to show the complexity of demand drivers to multiple audiences.

Conclusion

We argue that queue-length demand models poorly reflect systemic load. Our model allows discrete consumption measurement allowing for better resource management, problem solving and more agile health systems.

Keywords: Health, system dynamics, causal loop, demand predictors, demand modalities

4 - 2. Background

Josephine[15] had three years prior been diagnosed with Diabetes type II. One day, whilst walking barefoot, she cut her big toe on a small piece of glass. The wound was sore and relatively deep, but she tended to it by washing it out, and dressing it. With time, the inadequately treated foot introduced infection, which led to swelling and gangrene. When Josephine had three toes affected by gangrene, she went to the emergency department of the hospital near her workplace. After waiting for eight hours, it was discovered that the hospital was not in the district where she lived, and she was told to go to a different hospital. It took time before Josephine could return to the "correct" hospital. Again, considerable time had passed, and the gangrene had spread across her foot. Josephine was put on an emergency surgery waiting list and asked to return on a certain date for amputation. A few weeks later, Josephine returned and received a below the knee amputation of the lower limb. Josephine first spent six weeks in hospital for initial rehabilitation and received a prosthesis within 6 months. She then had to undergo a further year of rehabilitation, occupational therapy, and stump fit modifications.

The what ifs

What if, on the day she cut herself, she had gone to the clinic, had her lesion washed out, disinfected, dressed properly, and received antibiotics? Ten

---

[15] Not her real name. Josephine is a poor 43-year-old obese woman. She lives in informal housing and makes use of the public health care services in her developing country

minutes of nurse time, and five dollars of treatment, to alter the outcome completely.

What if on the day she first chose to go to the hospital, she had not been turned away from treatment, either because she had known to go to the correct hospital, or because hospital policy did not divert cases according to residence? In this case, Josephine would not have lost the time in going from one hospital to another. Even so, an amputation might already have been necessary, probably requiring the removal of several toes. This would have required the services of a surgical team, considerable nursing interventions, a short period of rehabilitation and a restoration to comparatively normal life within six months. In this scenario, the total cost would have been several thousand dollars.

What if once booked for amputation she did not have go onto a waiting list, but treatment was immediately available? In this case, a lower limb amputation would have reduced to a full - foot or perhaps even just a mid-foot amputation.

The story of Josephine tells an interesting, and for many developing countries, a common, story. The phenomena are very easy to understand for Diabetes-care but are mirrored in most diseases, only perhaps not quite so dramatically.

4 - 3. Structure of this Study

4 - 3.1. Context

This paper is the first in a three-part suite of papers that is concerned with understanding demand in health care. Figure 1-1 shows the composition of the whole body of work. Phase one, the focus of this paper, attempts to understand the units for measuring demand and how demand is constituted in health care. Phase two considers one of the building blocks from phase one – failure demand. phase two produces a mental model and an algorithm for identifying failure demand in a complex context. Phase three is an empirical study conducted in a national pharmaceutical supply-chain and presents the systemic impact of failure demand.

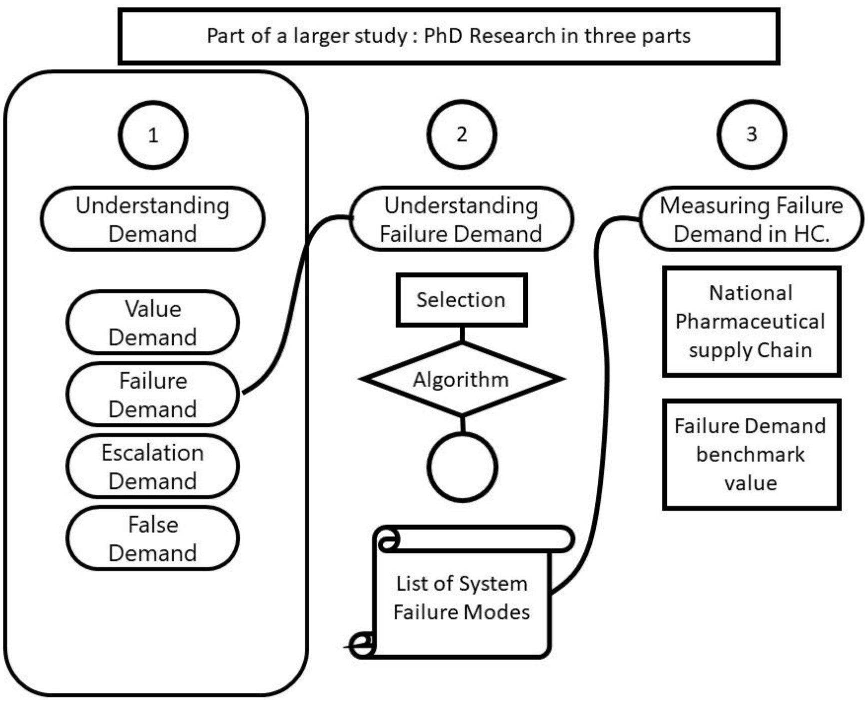


Figure 4-16: Context diagram for the suite of three publications.

4 - 3.2. Aims

A few authors have made attempts to understand the building blocks of demand in service industries, including tourism in Turkey [296], microcredit in Nigeria [297], private tutoring in Korea [298] and antenatal care in Colombia [299]. The question of why people are in a queue for health services, however, has been given surprisingly little attention over the past five decades.

The aims of this study are to:

- describe the demand population for health systems,
- describe the building blocks that shape demand in health care,
- propose a complex unit of measure for demand,
- represent the findings of this paper in a causal loop diagram

In other words, the framework, that is the product of this work, can be used to describe the underlying mechanisms that cause demand in the health system. The purpose of this study was to provide a framework of demand, not to solve it, hence, the system dynamic model [99] that is shown in Figure 4-19: is the ultimate output for this study, and can guide future research [300].

4 - 3.3. Method

To achieve its purpose, this study was structured as an exploratory-conceptual study [16], which aims to generate new knowledge in the form of a grounded framework, or a tool, that can be used in future studies [301], guiding research [300]. Data sources included literature sources, and our own observations and expertise from observations, experience and pilot studies [300].

Over the last ten years, we have conducted about a dozen, mostly analytical, studies and consulting projects, looking at the operation of the health system. This has revealed several demand phenomena which require exploration. Our observations were augmented by literature sources, that explore demand phenomena, to construct a robust model.

To establish our conceptual model, we made use of thought experiments [265, 267, 302], heeding the cautions to not descend into the absurd [303]. Our approach to understand the demand population and its origins was to ask questions such as: “what if sick people aren’t sick?”, “what if sick people don’t go to hospitals?” and similar questions. We then created a conceptual demand population model, the first model of this paper. After that, we built a conceptual model that defined demand categories, in which case, the use of the literature sources [19, 44, 232-234] were synthesised.

We validate our models through argumentation and vignettes, an approach justified if the narrative can clearly illustrate the logical progression, of thought, represents relatable, real scenarios and does not seek to manipulate the reader [304].

Our observations highlighted that demand is more complex than just heads in a queue, for which we designed a mental model of demand complexity which is aligned the thinking on the complexity in health care [176, 177].

4 - 4. Literature Review

This review presents the current thinking about health system demand, notably the demand population, the units of demand and existing models that describe the constituents of demand. The current models are synthesised, augmented by our own observations and experience, other studies, and published work. The novel elements of this paper, namely the units for complex demand and a demand modularisation are presented

Where goods are concrete, demand is easy to understand. In services however, demand is less clear [112]. Service demand is distinct from demand in other industries. Services cannot be buffered or stored [112], there is a high need for producer-consumer interaction, the consumer generally has to travel to the location of the service delivery and due to the intangible nature of the product, measuring capacity is often subjective and difficult [112] .

4 - 4.1. The demand population

Figure 4-17 is a conceptual model that shows a summary of the literature augmented by our own observations of how demand is structured in health care. The entirety of Figure 4-17 represents society, showing a finite, though for many practical purposes, unlimited underlying population. Society is split into a healthy population and an unhealthy population. To define what "unhealthy" means, we inverted the WHO definition for "being in good health"[16]. "Healthy" and "unhealthy" are dynamic sets, that vary in size and

[16] The WHO defines "being healthy" as "...a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" [305 Pg. 1] . We invert this to create a definition of "being

membership, as individuals get sick, and heal, with people moving from one to the other continuously. The shaded circle represents the portion of the underlying demand population that places a load on the health care system. In other words, the shaded circle represents the demand that the health system "sees".

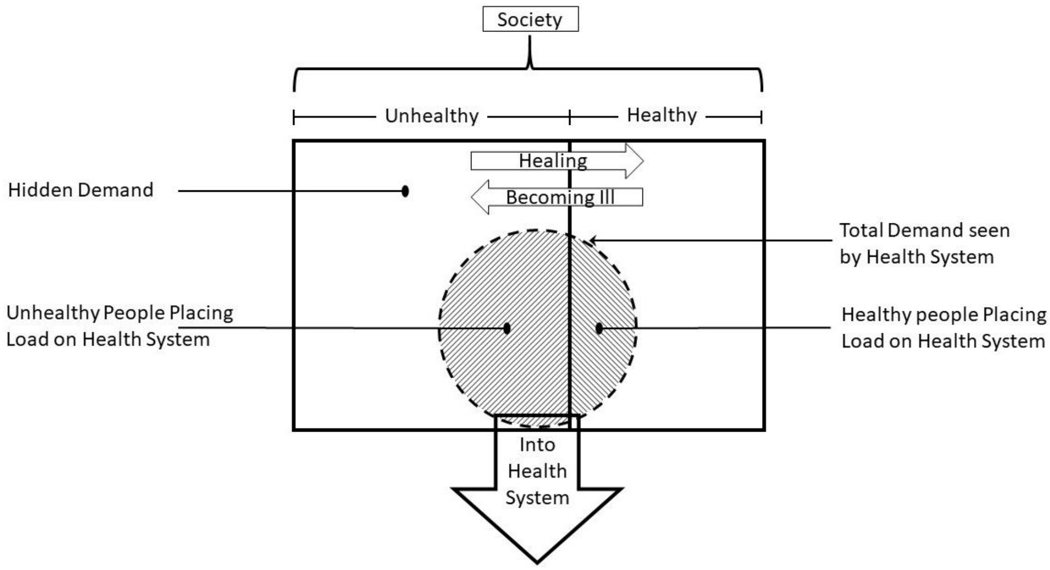


Figure 4-17: The demand population and the load on the health system

The demand experienced by the health system has multiple and unexpected origins. Firstly, it would be expected that the full unhealthy population is the demand population [104], but this is not the case. A proportion of unhealthy people do not seek health care [306]. The Mental Health services of NHS-Scotland (which we will refer to as "Mental Health Scotland") calls this portion of demand "hidden demand" [232]. This group is often de-motivated by cost [8], the inconvenience of pursuing treatment or the lack of a perceived need for care [19]. Hidden demand either resolves

---

unhealthy" as: "the absence of a state of complete physical, mental and social wellbeing or the presence of infirmity".

through the self-limiting nature of disease [307] or individuals choose to live in poor health, either intentionally or not. Hidden demand further represents a system risk, as its magnitude is by its nature uncertain and should a greater proportion of hidden demand place load on the health system, the additional burden could be significant.

Similarly, it might be expected that the entire healthy population falls outside the demand population of the health system, but again, this is not the case. For a variety of reasons healthy people can consume health resources, which includes administrative purposes, such as getting sick notes [19] or people who believe they are ill but may not be [308]. This modality of demand has been referred to as "inappropriate demand" by Rogers et al. [309], or "non-health motives" by Vickery & Lynch [19].

The composition of the shaded circle is of greatest interest for this study. This represents the work that the health system must do, and the greater this circle is, the higher the load on the system and the higher the consumption of system resources is [310]. We make a contribution to understanding how the shaded circle is structured.

4 - 4.2. Units of demand

This section explores the current thinking around the first major contribution of this paper, how demand is measured.

Demand in health care is frequently reduced to headcounts or waiting times. This is a common way for measuring demand [311]. This approach is used in the improvement literature, such as the works by Graban [13] and Toussaint [197] and even for health systems reporting by the World Health

Organization [312]. Seddon, however, criticises those who view service demand in terms of "units of production" as one would for manufactured goods. This can be reduced to standard processes and standard times can be controlled and enforced [44 Pg. 38]

This reductionist "units-of-production" [44] thinking assumes that all agents in a demand system place an equal load on the system [311], ignoring the impact of complexity, and variation [44] which is surprising, given that health care is largely a complexity driven domain [176] with considerable variation [210].

Among authors who have published work in "demand management in health care", the general definition for what demand is, gravitated towards the model of headcount in queues [313-317] or even only the number of admitted cases [318]. Broadly, the complexity of care [319] and the nature of demand in the service industry [112, 122, 320] are neglected when considering the nature of demand. Goldberg did mention the complexity of work as an important factor when trying to understand demand [319], though did not propose how such complexity might be structured.

This poverty of interpreting demand as anything other than queue length, is in spite of the fact that Grossman's seminal models from the Seventies [310], and further work by Wagstaff [321] propose the notion of the "stock of health capital" which supports the idea that demand is about the consumption of resources.

A good analogy to understand this, is electrical utilities that measure demand in Kilowatt hours, not in the number of devices plugged into the grid. What matters to the utility is how much of its product gets consumed, not the number of users [322]. Thus, a thousand mobile phone chargers may consume the same as one large oven. To the utility, this represents the same demand, and their production is based on consumption and not the number of devices.

Moving this analogy into the health sector, consider two queues of ten patients, the one queue consists of patients needing an aspirin, whilst the other queue has ten cancer patients. Although the demand interpreted on head count is the same, clearly the burden on the system differs.

The diverse views of demand in health care range from simplistic but practical models [152, 233] to more complex, but perhaps less practical models [310, 323]. Rather than presuming levelled demand based on the number of individuals, we will present our contribution regarding a novel approach to understanding complexity in demand. This will address the gap introduced between the realisation of complex demand and the consumption of health resources.

We believe that the use of time is a good proxy for the consumption of system resources, this has to be balanced against the level at which such resources are consumed.

4 - 4.3. Existing demand modularisation models

This section explores the second major contribution of this paper – the building blocks of demand.

Several theorists have proposed models to describe the building blocks of demand. ranging from those that largely conflate demand and the unhealthy population [104], to logistic regression models that identify race, location, sex and age as important demand shapers [306]. This section deals with those authors who have proposes a structured "arithmetic" model that shows the building blocks of demand for health care.

4 - 4.3.1. **The clinician's model**

Vickery & Lynch propose a model of demand in health care which addresses the question from the clinical point of view [19]. They classify demand into a relational model of four key categories, morbidity, perceived need, patient-preference and non-health motives. This deconstruction of demand is the most comprehensively argued modularisation currently in the literature. This relationship is represented in Equation 8:

$$D_t = D_m + D_d + D_p + D_n \quad (8)$$

With: D_t: Total demand; D_m: Morbidity demand; D_d: Perceived need; D_p: Patient preference and D_n: Non-health motives

The names of the modalities were very well chosen. Morbidity is the load of the system by those who certainly need care, the ill, the injured, the suffering [19]. Demand from perceived need is that load placed on the health system due to patients' opinion that they require care [19]. This is significant as it was found that people who perceive themselves to be in poor health put a greater load on the health care system [324]. Patient preference is strongly associated with informed consent and works in most instances as a moderating influence on load because patients tend to be more risk averse when better informed [325] with physicians themselves the most risk averse for their own treatments [326]. Finally, non-health motives (also referred to as "moral hazard" [327]) covers the unnecessary load placed on the health system [19] this includes sick leave, disability and workers' compensation assessments [19, 327].

Recalling Josephine's example: At all points along her story, she classified as "morbidity demand" [19]. In other words, she was meant to be at the hospital. Yet, the way in which her demand manifested, differed throughout this journey. Her motives, at all points were to legitimately seek care, however, the systemic transformation at the various points differed [96]. Vickery & Lynch' thinking focuses on motive – exploring patients' intentions and reasoning for seeking care. Vickery & Lynch' demand model is a clinically useful model as it answers the clinician's question of "how much of the demand that we see actually needs my input?" or "why will the next patient come through the door", and perhaps "how much of the demand can be channelled to non-clinical colleagues?".

The key weakness of the Vickery & Lynch model is that it does not differentiate between first time morbidity demand and repeating morbidity demand that was created in the system by not doing something, or doing something wrong for the patient [44]. Excluding failure demand means that by their model, even if a patient returns because a mistake was previously made, the demand remains morbidity demand i.e. the demand that they want to see [44].

4 - 4.3.2. **The system thinker's models**

John Seddon's thinking about demand proposes that demand always consists of two types of demand [44]. value demand; the demand that we want to see, and failure demand, the demand that comes about as a result of "not doing something or doing something wrong for the customer" [44 Pg. 76]. This binary classification is shown in Equation 9.

$$D_t = D_v + D_f \tag{9}$$

With: D_t *: Total demand;* D_v *: Value Demand;* D_f *: Failure Demand*

Value demand is the demand that the system wishes to see. It is the source of revenue for organisations and is the reason businesses exist [44]. Failure demand (previously called "the demand that we do not want" [136]) represents a meaningful opportunity for demand management and systems improvement. Seddon claims that "removing [failure demand] has an enormous impact on the economics of a system" [44 Pg. 38]. It has been shown that failure demand accounted for up to 45 % of load in financial services [44] whilst in the police and telecommunications this number went even higher, up to 80 %.

Failure demand originates when something that should be done isn't done, or when something that is done, is done wrong [44] and examples of this abound in Josephine's case). This means that when this workload returns, the work needs to be done again [232]. The implications are clear, if failure demand is reduced or eliminated, the load on the system can be significantly to overwhelmingly reduced. Put another way, if failure demand is ‘occasions where the system did something wrong’, that means, if next time ‘the system does is right’, demand can be reduced.

There is however a problem with Seddon’s binary categorisation. Given that demand is one or the other, if demand is not value demand (for example a healthy person who believes they require care, but do not) but is also not failure demand, as the system had not previously “not done something, or done something wrong” [44 Pg. 37] then that leaves this demand case beyond description. For that reason, further categories are required to describe the rather more complex demand landscape that exists in health care, compared to call centres and financial services [44].

Mental Health Scotland [232-234] have developed a modularisation of demand that is in use to understand demand in their system, shown in Equation 7 . Their model consists of “actual demand”, “failure demand”, “created demand” and “hidden demand”.

“Actual demand” is the term in use by Mental Health Scotland and it is defined as “What [the] service exists for” [233 Pg. 6]. This definition is similar to Seddon’s definition for “value demand” [44] and we treat the terms as interchangeable. They further define failure demand as “work that has to be done again because it was not done correctly the first time” [233 Pg. 6]. This is not Seddon’s pure definition for failure demand [44], yet the use of the same terminology again implies equivalence. Created demand is defined as “demand [that] is created because of how [the] service responds” [233 Pg. 6]. An example for created demand that was provided was scheduling unnecessary tests, consultations or other instances where poor work planning and structuring leads to a greater load on the system [232, 233].

The final type of demand identified by Mental Health Scotland is hidden demand (recall Figure 4-17), which exists in the unhealthy population but that does not place a load on the health system.

$$D_t = D_A + D_F + D_C + D_H \tag{10}$$

With: D_t *: Total demand;* D_A *:Actual demand;* D_F *: Failure demand;* D_C *: Created demand and* D_H *:Hidden demand*

Mental Health Scotland further differentiates between "total demand" as described above, and "current demand". Current demand excludes hidden demand and therefore represents the load that the system experiences, whilst total demand includes the unknown risk due to demand that is not realised [232, 233].

The clinician's model presents the health planners with useful insights to understanding demand, however is limited in its ability to drive system design decisions [174] and leaves little room to identify modes of demand that can be reduced or managed. The system thinkers view proposed by Seddon has an implicit binary nature and does not consider that demand can be anything other than value demand or failure demand. Mental Health Scotland proposes some novel categories, however one category, "created demand" has problems. This paper synthesises the clinicians' view – particularly the thinking about demand originating from the healthy population, with the thinking from Seddon around value and failure demand and augment this binary classification with some observations from the other models and from our own experience.

4 - 5. Analysis, synthesis, and results.

4 - 5.1. Complexity – the units of demand

Leading on from the literature presented in Section 4 - 4.2, we take the view that the units of demand are expressed in their consumption of health resources, rather than headcount [13], or time taken [232]. We favour the notion of the consumption of resources and use this to be essentially equivalent to load (on the system). The easiest way to view this in a simple system is the time consumed and we reject the "heads in queue" model as misleading.

We do however propose that to measure demand with greater texture, demand must be assessed in terms of a compound measure that expresses the complex nature of the burden on the system. This should consider the complexity of work and the level at which this takes place, and the time taken for its completion. We represent this symbolically in Equation 11, showing that total demand or system load is a function of procedural complexity and the time required to execute this.

$$Demand = f(complexity, time) \tag{11}$$

If we recall Josephine's case, the demand case that she represented changed dramatically from the point when she represented ten minutes of nurse's time to the escalated case where she required dozens of hours of attention from surgeons, anaesthetists, theatre nurses, rehabilitation staff, prosthetists, and many others in the care pipeline. Yet in both instances, she was simply one head in a queue, or perhaps more interestingly considering the complexity, many heads in many queues.

We propose that demand is a composite that assesses the expertise required, as well as the level of familiarity that is required to deliver a service. Our model was inspired by the Cynefin framework [54, 61] and classical 2x2 matrices like the Boston Consulting Group (BCG) – matrix [59]. In this case, we modify Snowden's "axes" of "abstraction" and of "learning to teaching" [54] and interpret those as "specialisation" and "familiarity". This builds on the notions of work complexity [319]. The matrix shown in Figure 4-18: proposes a structure with complexity rising clockwise towards the top right. The examples shown in the matrix demonstrate the usefulness of categorising demand events.

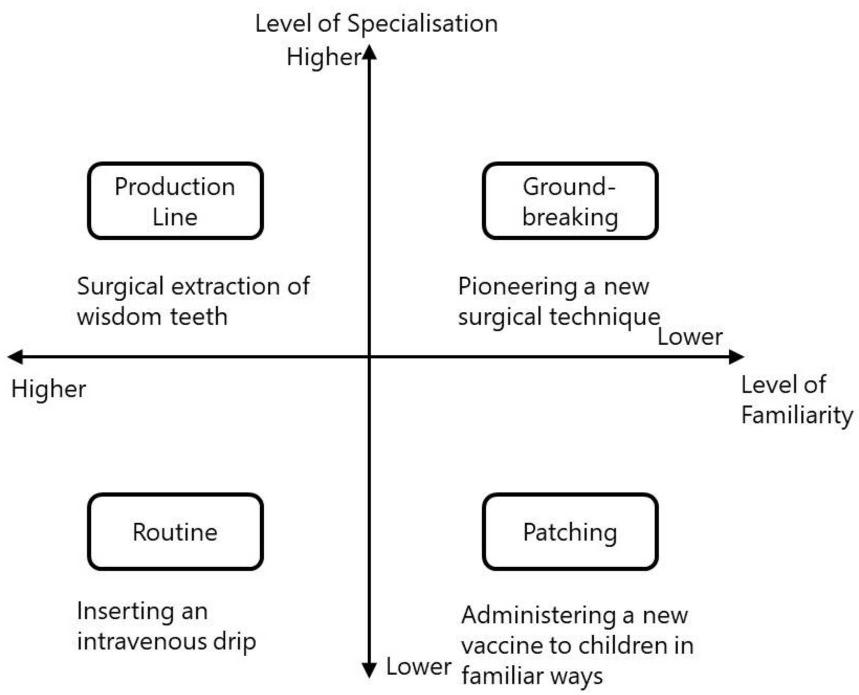


Figure 4-18: Conceptual complexity matrix

The actual values for the complexity multiplier will be determined by the utilisation of resources and their relative scarcity depending on the site. This requires further development. Based on this matrix, the highest demand multiplier therefore would be for a super-specialised innovative work packet. Of course, true demand complexity arises where, as for Josephine several work elements are required from many different service providers. It then becomes the sum of all complex interactions as shown in Equation 12:

$$Demand = \sum_{i=1}^{n} (complexity_i \times time_i) \tag{12}$$

With i as the numbered interaction and n being the total number of interactions.

Equation 12 builds on Equation 11 – meaning that demand is the complexity shown in Figure 4-18: multiplied by the amount of time such complex work consumes. This is intuitive, as ten minutes of super specialist time represents a very different demand case to ten hours of the same complexity of work. Similarly, ten hours of administration represents a different system load to ten hours of neurosurgery. This validatory vignette shows the utility of the tool and future work can test it in greater depth.

4 - 5.2. Novel modularisation

This section leads on from the context described in Section 4 - 4.3. The model by Vickery & Lynch makes contributes to understanding the amount of load that a health system is likely to encounter. It does however not differentiate between demand that is "new" and demand that is "repeated".

Seddon's binary model proposes that demand is either value demand, the demand that we want to see – or failure demand, the demand that we do not want to see. This binary – either-or-classification is difficult to translate into a service as complex as health care. Therefore, the addition of further terms is useful. Mental Health Scotland proposes a good model for their internal classification of demand. The addition of "hidden demand" gives a name to the demand that exists but does not place load on the health system.

We propose a demand model that includes four elements. This modularisation is shown in Equation 13. that shows how system effects can shape the load on the system.

$$D_t = D_v + D_e + D_f + D_n \tag{13}$$

With: D_t: total demand, D_v: value demand, D_e: escalation demand, D_f: failure demand, D_n: false demand

We use Seddon's terminology of value demand and failure demand. Where the load on the system, i.e., the demand increases due to a delay in treatment, we introduce the terminology escalation demand. We further make use of the terminology false demand – the demand that emerges from the healthy population.

4 - 5.2.1. **Value demand**

We use the terminology of Seddon [136] to refer to value demand, the true demand, the demand that we want to see, "the demand for which the service centre exists" [44 Pg. 37]. Value demand is the demand that customers place on the system, and value is determined by the point of view of the customer [122, 137]. Value demand arrives at the right time at the right place, with the right condition. In an ideal system, value demand is the only form of demand that should exist and all other forms of demand diminish the capacity of the system [40]. The three major non-value-demand modes are presented in the following sections.

4 - 5.2.2. **Escalation demand**

In Josephine's case, the resources that her condition consumed escalated over the period of delayed treatment. Some studies have explored the effect of delayed treatment, including Myocardial Infarction [328, 329] and Tuberculosis [330, 331]. In a major US study, it was found that one in eight delayed treatments leads to a condition becoming more severe [8]. The focus of these studies however was not to show how delayed treatment raises the consumption of health system resources [310], but rather to show the clinical effects, i.e. how the conditions deteriorate.

In our experience and from our observations the emergency rooms in developing nations are case studies in escalation demand. Conditions abound that could have been treated with minimal complexity at an earlier stage. Our vignette about Josephine is no fiction, with only specifics fabricated. This vignette is representative of our observations in at least half a dozen countries.

We introduce escalation demand as a novel modality that raises the load on the health system through delayed intervention. escalation demand, in system dynamic language [99, 104], is a reinforcing loop which raises the eventual demand after treatment delays [99]. As can be seen from

Josephine's case, the drain on system resources is considerable, and usually non-linear. Escalation demand can originate from patients choosing not to seek care, from care being unaffordable or unavailable, the belief that a condition will go away by itself or previously discouraging engagements in seeking care (which would make escalation demand in such cases a special case of failure demand).

Escalation demand is a curiosity of health care because load in other settings typically does not escalate due to delay. For example, the amount of work required to approve a town planning document or set up a client's router does not get altered by delays in engaging the service.

Escalation demand is an important contribution, as neither the language, nor the concept have been written about before. Escalation demand is a modality of demand that justifies further investigation. It is distinct from failure demand, as it is concerned with the increasing *severity* of demand due to delay whilst failure demand is concerned with the repeating *instance* of demand due to system failure.

4 - 5.2.3. **Failure demand**

Failure demand refers to the demand created by a feedback loop triggered by prior system failure. From Josephine's case, we recognise failure demand in the instances when she had to return to care twice because the system was unable to resolve her concerns, the system should be interpreted to include Josephine as an agent, and her actions are also system failure.

Failure demand is a reinforcing loop, that raises demand through repeated work interactions.

There is a threefold loss of time to the system through failure demand. Firstly, the customer loses all the time that they spend in the system the first time. Secondly the time lost by the person delivering the incomplete service. The third category of lost time is the patients who were deprived of care because resources were consumed by a patient who was not to experience resolution.

4 - 5.2.4. **False demand**

We introduce the terminology “false demand” to describe demand that originates from the healthy population. Vickery & Lynch proposed two modalities that emerge from the healthy population, namely perceived need, and non-health motives [19]. As such they are absorbed into our broader category of “false demand”.

Examples of false demand includes people who attend a health setting to receive medical certificates or complete other administrative tasks [19], hypochondria [308] and even as in the case of an Irish study, lonely (elderly) people who seek companionship at hospitals [332].

4 - 5.2.5. **A note on abandoned modalities**

Two modalities that have been proposed by the model from Mental Health Scotland are excluded from our final model, the reasoning is given in this section.

Created demand

The definition of created demand emphasises that it originates from poor decision making or actions by the health care provider leading to additional load on the system. This is a very clear example of "not doing something right for the [patient]" [44 Pg. 37] and is as such a prototypical example of failure demand for which reason we conflate them in this paper.

Hidden demand

Hidden demand does not (currently) place a load on the health system (as can be seen in Figure 4-17), however the risk exists that it may. For that reason they differentiate between total demand (that includes the risk from hidden demand) and actual demand the current load on the health care system [232]. Although we agree with the idea of hidden demand and have included it in Figure 4-17, we exclude it from our final model as it only describes the impact of the load experienced by the system. Should hidden demand be transformed into load on the health system, then this load will take the form of one of the modalities we have presented in this paper.

4 - 5.3. System dynamic model

Systems thinking, causal loop diagrams and their computational extension into system dynamics, are useful approaches for representing complex mental models[17].

Causal loop diagrams show relationships between phenomena, primarily to show the effect of modifiable elements on system performance. System dynamics has been used to model large, complex strategic systems, which rely on "flow" [333]. As such, these techniques have been used in disease epidemiology [334, 335], patient flows [336], evaluating health care capacity

[17] For an excellent introduction to Causal loop diagrams in health settings, the reader may look at the work of Homer [104].

[314] and the interaction between these factors. We unify our demand model and our various modularisations for demand into the mental model shown in Figure 4-19: [18].

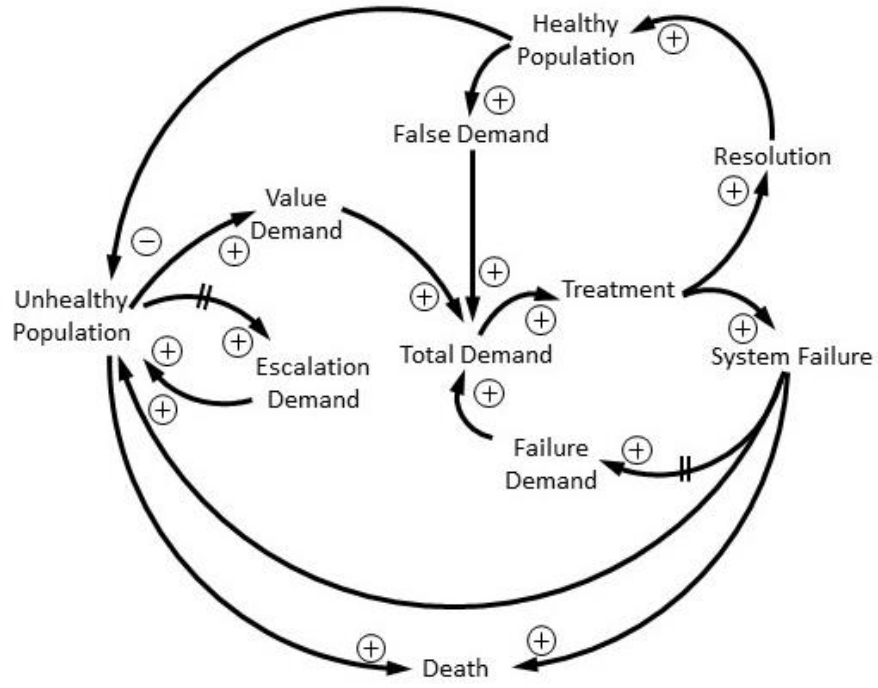


Figure 4-19: Causal loop diagram of the modularisations of demand in a health system[19]

[18] Causal loop diagrams represent relationships between phenomena. A note to understand the conventions. Lines that tie phenomena together indicate a relationship. An arrow with a + represents a positive relationship, meaning that trends match. For example, if two items are positively matched, if one goes up, the other does too, and importantly, if one goes down so does the other. A negative relationship, indicated by "-" means that when the one phenomenon goes up, the other phenomenon goes down and vice versa. Parallel lines cutting a relationship line indicate a time delay

[19] The feedback flow of stock from the unhealthy population to the healthy population as described in section 4 - 4.1 is omitted for clarity.

4 - 6. Discussion

One reason why so many simplified metrics are used to assess demand (as in Section 4 - 4.2) may be because counting people in queues is easy, as is calculating process time or assessing waiting times. However, complex measures as we propose, are difficult to use and, in many instances, impractical to use on an ongoing basis. We propose that the complex measurement of demand is a more accurate approach, however we recognise that practically this is difficult and may be of little value for daily use. We propose, nevertheless, that using the complex measure would give a more accurate insight into demand which will be of use when strategic decisions are required [174]. builds on the figure seen earlier and summarises the thinking presented in this paper related to the building blocks of demand. It shows the modalities of demand as identified in Equation 13 but also captures hidden demand even though it has no current influence on the load experienced by the health system.

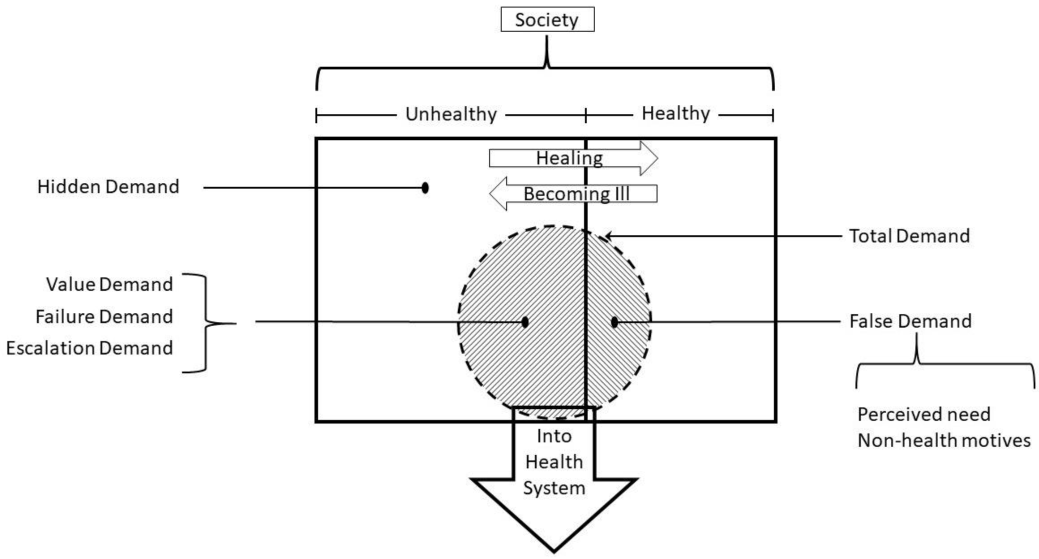


Figure 4-20: Summary of the demand model proposed in this paper

The model that we have proposed is conceptual, and the rates of failure and escalation demand should be determined empirically to process the system dynamic model computationally.

We recognise the risk posed to the health system by hidden demand, the need that exists that is not matched by a transformation into load on the health system. This must be considered when strategic system design is done or policy is set [174, 337]. There is no guarantee that hidden demand will remain "hidden", especially if barriers, such as cost [8, 338] are lifted.

The causal loop diagram shows two strong reinforcing loops which have the potential to dramatically drive total demand: – escalation demand and failure demand. Although managing these may not be easy, understanding these two phenomena better and designing systemic strategies to reduce their negative impact will be vital.

Key interventions will include awareness campaigns for medical staff as well as for patients. The negative impact of both modalities can be limited through awareness in all levels of the health care system, from the operational level up to the strategic level, which would be useful when designing health systems [174]. This would allow for policies to be designed to that consider the effect of these demand modalities.

4 - 7. Suggestions for further research

This study has contributed several novel mental models that can be of help to clinical managers, and district and national planners of health services. We propose that other researchers may be interested in conducting further research into:

- The complexity model needs to be calibrated and modified to reflect regional and contextual differences.
- The demand complexity model can be tested in a qualitative study using a method similar to that of Yu et al. [323].
- The demand complexity model needs to be used to build a set of benchmark values for common demand scenarios to build a more practical assessment tool, thus emerging from the theoretical/academic realm into
- The demand modalities may be measured in different contexts and then used as input values for the system dynamic model presented in this paper.
- Although a major contribution, testing escalation demand would require a complex qualitative study, interviewing patients, with a strong clinical team to assess the impact of escalation demand given that the end points may be concrete, but the start points would be obscure

4 - 8. Conclusion

We propose a population model that splits society into healthy and unhealthy people. To do so, we propose a definition for "unhealthy" people. We presented a model that tries to define the demand population for the health system. We present a complex model that defines demand by the extent to which the resources of the health system are consumed [121] is drained. We argue for the presence of value demand, which is true demand, presenting at the right place, at the right time, and that this must be the mission of a system to serve. We further show that non-value-demand depletes system resources with limited benefit to the system and patients. Failure demand is joined by novel modalities that we introduce, namely, escalation demand – the rising demand due to time delays – and false demand as originating in the healthy population. We present a causal loop diagram which is a mental model that explains the relationships that create demand on the health system and identify opportunities for improvement, which may include reduction of demand in the long term, thus freeing up health capacity.

4 - 9. References

8. Dijulio, B., Kirzinger, A., Wu, B., and Brodie, M., *Data Note: Americans' Challenges with Health Care Costs.* Kaiser Family Foundation, 2017.
13. Graban, M., *Lean Hospitals*. 2008, CRC Press, Boca Raton, London, New York
16. Yin, R.K., *Case Study Research: Design and Methods*. 1987, Sage Publications, Inc., Beverley Hills
19. Vickery, D.M. and Lynch, W.D., *Demand Management: Enabling Patients to Use Medical Care Appropriately.* Journal of Occupational and Environmental medicine, 1995. 37(5), p. 551-557.
40. Hopp, W.J. and Spearman, M.L., *Factory Physics*. 2011, Waveland Press, Long Grove, Illinois
44. Seddon, J., *Freedom from Command and Control: Rethinking Management for Lean Service*. 2005, CRC press, Boca Raton, Florida
54. Snowden, D.J. and Boone, M.E., *A Leader's Framework for Decision Making.* Harvard Business Review, 2007. 85(11), p. 68.
59. Madsen, D.O., *Not Dead Yet: The Rise, Fall and Persistence of the Bcg Matrix.* Problems and Perspectives in Management, 2017. 15(1), p. 19-34.
61. Hasan, H. and Kazlauskas, A., *The Cynefin Framework: Putting Complexity into Perspective.* Univerity of Wolongong, 2014.
96. Checkland, P., *Systems Thinking.* Rethinking Management Information Systems, 1999, p. 45-56.
99. Senge, P.M., *The Fifth Discipline.* Measuring Business Excellence, 1997. 1(3), p. 46-51.
104. Homer, J.B. and Hirsch, G.B., *System Dynamics Modeling for Public Health: Background and Opportunities.* American Journal of Public Health, 2006. 96(3), p. 452-458.
112. Sasser, W.E., *Match Supply and Demand in Service Industries.* Harvard Business Review, 1976. 54(6), p. 133-140.
121. Jones, D.T. and Womack, J.P., *Lean Consumption.* Harvard Business Review, 2005. 83, p. 58-68.
122. Womack, J., *Lean Consumption.* Manufacturing Engineer, 2005. 84(4), p. 5-5.
136. Seddon, J., *I Want You to Cheat!: The Unreasonable Guide to Service and Quality in Organisation*. 1 ed. 1992, Vanguard Press, Buckingham, UK

137. Seddon, J., O'donovan, B., and Zokaei, K., *Rethinking Lean Service*, in *Service Design and Delivery*. 2011, Springer. p. 41-60.
152. Hartmann, D. and Mandavha, R. *Lean Healthcare, a Casualty of Inefficiency*. in *SAIIE Conference*. 2010. Pretoria, Southern African Institute for Industrial Engineering.
174. Hines, P., Holweg, M., and Rich, N., *Learning to Evolve: A Review of Contemporary Lean Thinking.* International Journal of Operations & Production Management, 2004. 24(10), p. 994-1011.
176. Plsek, P.E. and Greenhalgh, T., *Complexity Science: The Challenge of Complexity in Health Care.* BMJ: British Medical Journal, 2001. 323(7313), p. 625.
177. Plsek, P.E. and Wilson, T., *Complexity, Leadership, and Management in Healthcare Organisations.* BMJ: British Medical Journal, 2001. 323(7315), p. 746.
197. Toussaint, J. and Gerard, R., *On the Mend: Revolutionizing Healthcare to Save Lives and Transform the Industry*. 2010, Lean Enterprise Institute, Cambridge, Massachusetts
210. Jack, E.P. and Powers, T.L., *A Review and Synthesis of Demand Management, Capacity Management and Performance in Health-Care Services.* International Journal of Management Reviews, 2009. 11(2), p. 149-174.
232. Donnelley, R., *Mental Health Collaborative Programme Toolkit*, T.S. Government, Editor. 2009, Crown, Edinburgh.
233. Healthcare Improvement Scotland, *A Guide to Understanding and Managing Demand*. 2019, Mental Health Access Improvement Support Team (MHAIST), Edinburgh, UK.
234. El-Ghorr, A., Cameron, R., Fleming, M., Mckechnie, L., Thomson, D., and Doherty, S., *Scotland's National Approach to Improving Mental Health Services: Integrated Care Pathways as Tools for Redesign and Continuous Quality Improvement.* 2010. 14(2), p. 57-64.
265. Sorensen, R.A., *Thought Experiments*. 1998, Oxford University Press on Demand, Oxford, UK
267. Brown, J.R., *Review of Thought Experiments by Roy Sorensen.* Canadian Journal of Philosophy, 1995. 25(1), p. 135-142.
296. Uysal, M., *The Determinants of Tourism Demand*. The Economic Geography of the Tourist Industry: A Supply-Side Analysis. Vol. 79. 1998

297. Balogun, O. and Yusuf, S., *Determinants of Demand for Microcredit among the Rural Households in South-Western States, Nigeria.* Journal of Agriculture and Social Sciences - Pakistan, 2011.
298. Kim, J.-H. and Park, D., *The Determinants of Demand for Private Tutoring in South Korea.* Asia Pacific Education Review, 2010. 11(3), p. 411-421.
299. Vecino-Ortiz, A.I., *Determinants of Demand for Antenatal Care in Colombia.* Health Policy, 2008. 86(2-3), p. 363-372.
300. Ravitch, S.M. and Riggan, M., *Reason & Rigor: How Conceptual Frameworks Guide Research*. 2016, Sage Publications, Newbury Park, California
301. Collier, D. and Levitsky, S., *Democracy with Adjectives: Conceptual Innovation in Comparative Research.* World Politics, 2011. 49(3), p. 430-451.
302. Häggqvist, S., *A Model for Thought Experiments.* Canadian Journal of Philosophy, 2009. 39(1), p. 55-76.
303. Gendler, T.S., *Thought Experiment: On the Powers and Limits of Imaginary Cases*. 2014, Routledge, Abingdon, UK
304. Stoves, J., Connolly, J., Cheung, C.K., Grange, A., Rhodes, P., O'donoghue, D., et al., *Electronic Consultation as an Alternative to Hospital Referral for Patients with Chronic Kidney Disease: A Novel Application for Networked Electronic Health Records to Improve the Accessibility and Efficiency of Healthcare.* Qual Saf Health Care, 2010. 19(5), p. e54-e54.
305. World Health Organization, *Constitution*. 1989, World Health Organization, Geneva, Switzerland
306. Havemann, R. and Van Der Berg, S., *The Demand for Health Care in South Africa.* Journal for the Study of Economics and Econometrics, 2003. 27(3), p. 1-27.
307. Moerman, D.E., *Meaning, Medicine, and the" Placebo Effect"*. Vol. 28. 2002, Cambridge University Press Cambridge
308. Borossa, J. and Rooney, C., *Fortress Hypochondria: Health and Safety*, in *Nationalism and the Body Politic*. 2018, Routledge. p. 5-19.
309. Rogers, A., Hassell, K., and Nicolaas, G., *Demanding Patients?: Analysing the Use of Primary Care*. 1999, Open University Press, Maidenhead, UK
310. Grossman, M., *On the Concept of Health Capital and the Demand for Health.* Journal of Political Economy, 1972. 80(2), p. 223-255.

311. Stevens, B.H. and Rydell, C.P., *Spatial Demand Theory and Monopoly Price Policy.* Papers in Regional Science, 1966. 17(1), p. 195-204.
312. World Health Organization, *World Health Statistics 2014.* 2014.
313. Almeida, A.S. and Cima, J.F., *Demand Uncertainty and Hospital Costs: An Application to Portuguese Public Hospitals.* The European Journal of Health Economics, 2015. 16(1), p. 35-45.
314. Brailsford, S.C., Lattimer, V., Tarnaras, P., and Turnbull, J., *Emergency and on-Demand Health Care: Modelling a Large Complex System.* Journal of the Operational Research Society, 2004, p. 34-42.
315. Chapple, A. and Rogers, A., *'Self-Care' and Its Relevance to Developing Demand Management Strategies: A Review of Qualitative Research.* Health & Social Care in the Community, 1999. 7(6), p. 445-454.
316. Gallagher, M., Pearson, P., Drinkwater, C., and Guy, J., *Managing Patient Demand: A Qualitative Study of Appointment Making in General Practice.* The British Journal of General Practice, 2001. 51(465), p. 280-285.
317. Gillam, S. and Pencheon, D., *Managing Demand in General Practice.* BMJ : British Medical Journal, 1998. 316(7148), p. 1895-1898.
318. Fries, J.F., Koop, C.E., Beadle, C.E., Cooper, P.P., England, M.J., Greaves, R.F., et al., *Reducing Health Care Costs by Reducing the Need and Demand for Medical Services.* New England Journal of Medicine, 1993. 329(5), p. 321-325.
319. Goldberg, S.E., *Demand Management: Implementing Your Own Program.* Fam Pract Manag, 1998. 5(8), p. 49-50, 55-6, 59-62.
320. Seddon, J. and Caulkin, S., *Systems Thinking, Lean Production and Action Learning.* Action Learning: Research and Practice, 2007. 4(1), p. 9-24.
321. Wagstaff, A., *The Demand for Health: Some New Empirical Evidence.* Journal of Health Economics, 1986. 5(3), p. 195-233.
322. Brennan, T.J., *Demand-Side Management Programs under Retail Electricity Competition*. 1998, Resources for the Future Washington, DC
323. Yu, D., Lowndes, B., Thiels, C., Bingener, J., Abdelrahman, A., Lyons, R., et al., *Quantifying Intraoperative Workloads across the Surgical Team Roles: Room for Better Balance?* World Journal of Surgery, 2016. 40(7), p. 1565-1574.

324. Connelly, J.E., Philbrick, J.T., Smith Jr, G.R., Kaiser, D.L., and Wymer, A., *Health Perceptions of Primary Care Patients and the Influence on Health Care Utilization.* Medical care, 1989, p. S99-S109.
325. Wennberg, J.E., *Outcomes Research, Cost Containment, and the Fear of Health Care Rationing*. 1990, Massachusetts Medical Society, Waltham, Massachusetts.
326. Domenighetti, G., Casabianca, A., Gutzwiller, F., and Martinoli, S., *Revisiting the Most Informed Consumer of Surgical Services: The Physician-Patient.* International Journal of Technology Assessment in Health Care, 1993. 9(4), p. 505-513.
327. Donaldson, C. and Gerard, K., *Countering Moral Hazard in Public and Private Health Care Systems: A Review of Recent Evidence.* Journal of Social Policy, 1989. 18(2), p. 235-251.
328. Horne, R., James, D., Petrie, K., Weinman, J., and Vincent, R., *Patients' Interpretation of Symptoms as a Cause of Delay in Reaching Hospital During Acute Myocardial Infarction.* Heart, 2000. 83(4), p. 388-393.
329. Leslie, W., Urie, A., Hooper, J., and Morrison, C., *Delay in Calling for Help During Myocardial Infarction: Reasons for the Delay and Subsequent Pattern of Accessing Care.* Heart, 2000. 84(2), p. 137-141.
330. Storla, D.G., Yimer, S., and Bjune, G.A., *A Systematic Review of Delay in the Diagnosis and Treatment of Tuberculosis.* BMC Public Health, 2008. 8(1), p. 15.
331. Van Der Werf, M., Chechulin, Y., Yegorova, O., Marcinuk, T., Stopolyanskiy, A., Voloschuk, V., et al., *Health Care Seeking Behaviour for Tuberculosis Symptoms in Kiev City, Ukraine.* The International Journal of Tuberculosis and Lung Disease, 2006. 10(4), p. 390-395.
332. Molloy, G.J., Mcgee, H.M., O'neill, D., and Conroy, R.M., *Loneliness and Emergency and Planned Hospitalizations in a Community Sample of Older Adults.* Journal of the American Geriatrics Society, 2010. 58(8), p. 1538-1541.
333. Brailsford, S. and Hilton, N., *A Comparison of Discrete Event Simulation and System Dynamics for Modelling Health Care Systems.* 2001.
334. Roberts, C. and Dangerfield, B., *Modelling the Epidemiological Consequences of Hiv Infection and Aids: A Contribution from*

Operational Research. Journal of the Operational Research Society, 1990. 41(4), p. 273-289.
335. Rauner, M. *Managing the Aids Epidemic in Vienna, Austria: Prevention Strategies for the 21st Century*. in *Information, Management and Planning of Health Services, Proceedings of the 25th Meeting of the European Working Group on Operational Research Applied to Health Services*. 1999.
336. Davies, R. and Davies, H., *Modelling Patient Flows and Resource Provision in Health Systems.* Omega, 1994. 22(2), p. 123-131.
337. Bowen, S. and Zwi, A.B., *Pathways to "Evidence-Informed" Policy and Practice: A Framework for Action.* PLoS Med, 2005. 2(7).
338. Ahmed, S., Begum, T., and Cotlear, D., *Bangladesh: Unravelling the 'Good Health at Low Cost' Story*. 2019, World Bank Group, Washington D.C.

Chapter Five: **Understanding Failure Demand in Health Care: A Mental Model for Demand Management**

5 - 1. Abstract

The load on health systems due to systemic overburden, leads to heightened cost, longer waiting times, reduced quality of care, and associated problems. This may be caused by "failure demand" however its definition is inadequate for a complex hierarchical system. Although accounting for a significant proportion of load in other industries, the academic assessment of failure demand in health care remains limited. We present a novel way of identifying repeat consumption, which we loosely equate with failure demand. We present a framework that can be used to identify "system failure", the trigger for later repeat consumption. This provides new insight to understand whether common events represent system failure. A diagnostic framework was developed from observations, literature, and brainstorming. Commonly observed exit-scenarios in health care were tested against the framework to create a system-failure list. The framework and the categorisation table were shared with eight international Lean health care experts. Following feedback, the framework and categorisations were fine-tuned, and consensus achieved via member-checking. Identifying, and managing failure demand for these settings can lead to reduced system-load thus reducing cost, increasing system efficiency and quality.

Key Words: Repeat Consumption, System Failure, Failure demand; Health care, Efficiency, Demand Management, Quality, Health Policy

5 - 2. Introduction

5 - 2.1. Study context

The genesis of this study lies in trying to understand the phenomenon of failure demand, how it presents in health systems and the impact that this has on service delivery.

The first publication in a series of three publications (shown in Figure 5-21) identified five demand modalities in health systems. Of which failure demand was one. Recognising that there are gaps in defining and identifying failure demand in more complex hierarchical organisations, greater depth of investigation was required.

This paper forms the second part of a larger study that was conducted with the intention to understand certain aspects of demand in health systems. In this paper we present a framework that can be used to assess system failure – that could lead to failure demand. Common events in the provision of health care were tested against this algorithm, to validate it. We summarise our findings with a list of events that could be root causes of failure demand. The totality – framework and findings – is expert validated.

In the third publication, one of the categories responsible for failure demand identified in this paper, namely poor supply-chain management was explored in greater depth in an empirical study in a national pharmaceutical supply-chain in a developing country.

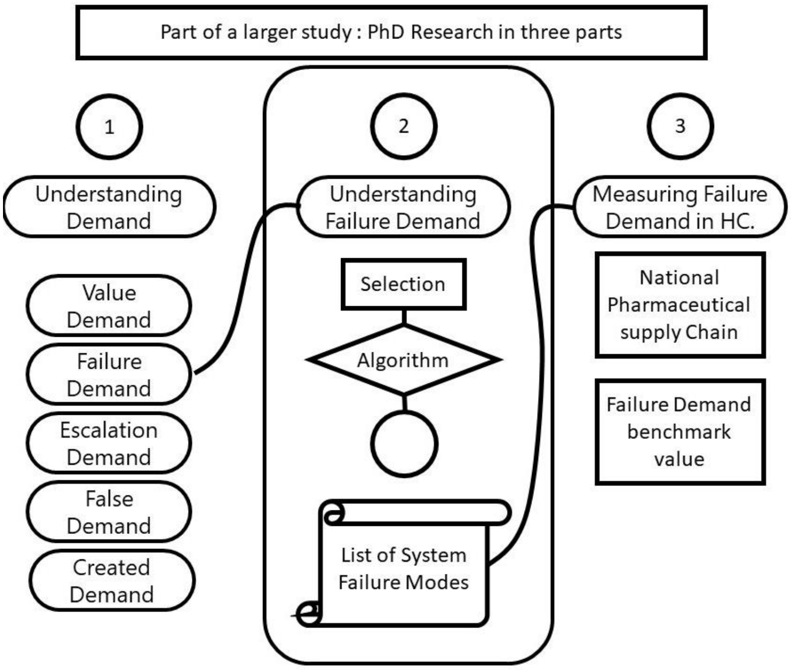


Figure 5-21: Context of this paper

5 - 2.2. Seeing failure demand everywhere

The Institutional Review Board (also referred to as Ethics- [review] -board or -committee[20]) at the authors' home institution recently circulated the following memo:

"We are aware of the concern of applicants about the length of time to receive responses to their applications submitted through the Research Office. The reason is that the committee and office are overwhelmed by workload. A recent audit shows that 891 new applications were submitted for review in 2016. Only 107 (12%) were approved at the first evaluation, 784 (88%) had to be resubmitted – this means that the new application workload was 891 + 784 = 1675 over and above other work. In 2008

[20] Nomenclature differs across the world; the North American standard form is "Institutional Review Board" whilst the British (and broadly commonwealth) naming is some variation that contains the word "Ethics" [339]. The functions of these are functionally indistinguishable and are governed by the declaration of Helsinki [340]

217/586 (37%) were approved at first evaluation after which there has been a steady deterioration. We apologize for the delays that are influenced by the workload."

This memo reveals a fascinating phenomenon that raises workload, increases demand on limited resources and due to this, increases waiting time and affects quality of work. Failure demand is the customer-interaction that occurs more than once because a previous interaction with the system that provides the service was unsuccessful.

5 - 2.3. Failure demand and resolution

Seddon [136] proposes two forms of demand: *value demand*, which (in the above example) can be thought of as the initially successful 12 % of ethics applications, and failure demand, the remainder of cases, which need the work elements to be repeated with subsequent reprocessing. He introduced the idea [44] as an evolution of his earlier thinking on "demand that we do not want" [136]. The definition of failure demand is:

> *"demand caused by a failure to do something or do something right for the customer". [44 Pg. 27]*

The remainder of this section unpacks Seddon's wording to visualise failure demand and identify its nuances. Consider Figure 5-22 that shows the current thinking.

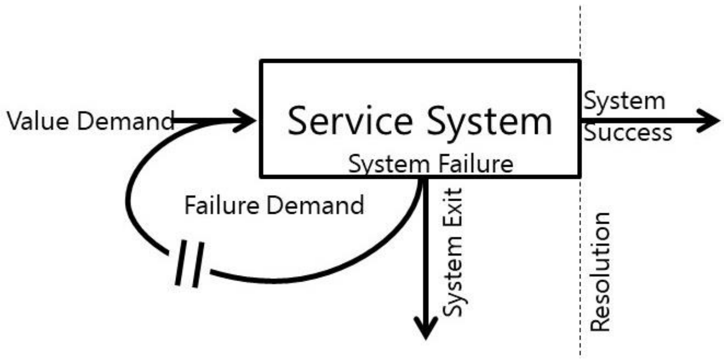


Figure 5-22: Generic system, showing system failure, failure demand and value demand.

The system receives a certain demand which consists of value demand and failure demand. This total demand is addressed by the system. A certain proportion of work passes through the system unimpeded, and that we refer to as system success. This is the normal case as proposed as the "conservation of material – law" by Hopp & Spearman [40] and the flow through system that forms the basis for Little's law [47].

5 - 2.4. What is resolution?

The notion of resolution is complex. Whereas in manufacturing it is clear when the product is delivered at its nominal value [131], in service systems, whether the nominal value of the service has been achieved is decided by the customer. This idea is implied in the service-directed work by Womack & Jones [121] and explicitly stated by Seddon [44]. Whether a service has been delivered to completion is therefore dynamic and no static value defining an acceptable "product" exists [44]. Attempting to standardise service offerings escalates cost without generally achieving greater levels of customer satisfaction [134]. This is further complicated in health care, as the objective of work often differs. Lillrank et al. differentiate between the purpose of treatment being either “Cure” where the target outcome is healing, whilst “Care” by contrast does not target a healed state, but rather focuses on maintenance [241]. For this reason, patients may need to return for treatment repeatedly without it being thought of as the system failing. If resolution was achieved. Analytically, this means, that failure demand cannot be measured by simply counting every instance of a patient returning (for the same reason).

Some work elements do not successfully progress through the system for a variety of reasons, and at various stages of the process. We call the trigger for this unsuccessful process: system failure. Once system failure has occurred, the work element may depart the system forever, which we refer to as system exit or the work element may return to the demand queue at a later stage after a time delay. The work elements that undergo system failure, yet return, represent failure demand. This means, that the total demand queue is elongated by failure demand elements, which place the same burden on the system as they did previously.

Whilst the classical definition of failure demand may suffice in some industries, we view health care as an environment in which Seddon’s definition requires refinement.

In this paper, we will discuss failure demand broadly. We will identify where failure demand comes from, what causes it, and what effect it has on overall system performance. We will describe the moment when the system is unable to serve the customer and introduce the terminology system failure to describe this moment. We introduce the drivers for system failure and explore how this moment of system failure relates to the occurrence of failure demand. We will show that although system failure is always the trigger for failure demand, not all instances of system failure result in failure demand. We will then show how the current definitions for failure demand are not enough to describe health care systems, due to their higher levels of complexity, complicated hierarchical structure, and characteristics of health care, including independent patient behaviour. We will propose an algorithm that will be used to assess system exits prior to resolution. This algorithm will be used to establish whether a particular modality of system failure leads to failure demand and tabulate our findings.

5 - 2.5. Examples of failure demand.

Seddon [44] primarily focussed the introduction of his idea of failure demand in the call centre industry. An example can be found in the case of a UK bank where it appeared that demand was increasing because the call volumes were increasing, so the bank in turn increased the number of call centres to cater to this demand. However, the reality was that call volumes had increased due to failure demand in addition to value demand. Failure demand in this case was found to account for 46% of the total demand. The remedy was to manage first time interactions, and as a result, the number of calls reduced substantially in the long term, leading to the ability to reduce the number of call centres. These studies on call centres have been retested and explained in numerous research papers over the years, failure demand levels ranging from 12 % to 80 % [341]. Marr & Neely [185] found that most studied organisations spend at least half of their time dealing with failure demand. This is also shown in our Institutional Review Board processing ethics applications – where the bulk of time is spent on reprocessing applications, which led to a stagnation of actual throughput. This is predictable using Little's law [47] showing that more feedback loops into the system lower the throughput rate.

The implications of failure demand is that most organisations are either overcapitalised to deal with such unnecessary demand, or alternatively their service delivery is compromised, which Piercy & Rich [342] identified as major opportunities for improvement.

5 - 2.6. Introducing *system failure* as a driver for failure demand.

Recalling our example of the Institutional Review Board: when an application is rejected, the system has failed. We propose that the applicant forms part of this complex system, and that the agency for the "failure" may lie there. The trigger that caused the system exit is the point of system failure. Applicants may wish never to resubmit their application, in which case, the system failure was the last act of the interaction. If applicants resubmit their ethics documents, then by joining the queue and forming a repeated burden on the system, that application becomes failure demand, in other words, demand due to a prior failure of the system.

Failure demand is the consequence of an event and does not exist in isolation of systemic triggers. Understanding these triggers, allows for the measurement, and the understanding of failure demand, which may lead to increases in effectiveness and efficiency [343]. Although implied in Seddon's definition a specific event triggers failure demand, and that is the instance of the system failing to do something or failing to do something right [44].

These triggers have previously been described using the terminology "failure" or "service failure" [185]. We prefer to refer to these triggers as "system failure" (terminology used almost in passing, by Piercy & Rich [342]) to identify events that contribute to unsuccessful work completion. When a user engages a system, but does not get resolution, and departs the system, the system has failed to meet the user's needs. System failure becomes failure demand only when customers return to the system in need of the same service.

The relationship therefore between system failure and failure demand, is one of cause and effect. This paper builds on the traditional model, considering the idiosyncrasies of health care, presenting a framework for identifying system failure in health care, which, if the patients return, results in failure demand.

5 - 2.7. Unique characteristics of health care systems.

Health systems in the 21st century are complex organisms. Plsek & Greenhalgh [344] identify many elements that categorise health systems as complex adaptive systems, which have characteristics that are distinct from simple systems. This complexity stems from the nature of health care as a commodity, the interaction between health care and patients and the nature of disease.

Unlike manufactured products, services cannot buffer completed product for release when demand occurs, meaning that full-service offerings can only be started when they are needed [112]. Whilst Sasser wrote that service demand tends to be personal and in person [112], meaning that nobody can receive a health care service other than the patient, and usually the care provider cannot have a proxy.

Moreover, because health systems tend to be capacity led, little attention is paid to demand management [141]. Walley found that this was particularly true in public services. Describing them as "resource-driven", he concludes that service delivery could be meaningfully improved through the adoption of private sector inspired demand driven strategies [186].

When trying to understand, or improve a service system, demand must be investigated and understood first [44, 98, 186, 194]. Designing solutions that do not consider demand first are likely to result in incorrect solutions or solutions to incorrect problems [44, 98]. To understand demand, it is important to not only know how much demand a system experiences [112], but the frequency or distribution of its arrival, what type of demand it is and in which units the demand is measured.

The nature of disease makes health care an even more complex service environment. Disease can present in many ways, responses to treatment vary and mistakes are made [345]. Often medical practice, although conservative by nature [346], is experimental, and correct treatment regiments are decided upon through trial and error.

5 - 2.8. Understanding how demand is measured.

Although demand can be reduced to a simple measure such as "the number of people in a queue" which is a view that we have taken in some earlier work [152], and also in some of the classical Lean health care literature [13, 197], this approach does not recognise that individuals cannot be equated with the load on a system. In later work, we made use of time consumed as the measure of demand [207]. But this did not fully address the true load on the system, differentiating between levels of specialisation, resource scarcity and work complexity. Wagstaff introduces the idea of the "stock of health capital", the commodity that health systems invest in the wellbeing of the community. This is the resource that is being depleted when a load is placed on a system [198]. To understand how this "stock" is structured requires that demand be seen in terms of the underlying complexity of task (which guides the decision of which resources are required to perform the work) and the time that is required to complete the work.

Therefore, the view we take of demand in this paper, is a composite that considers the amount of time that resources are consumed, and the type of resources.

5 - 2.9. The relationship between demand and capacity.

Utilisation can broadly be defined as the proportion of capacity being used for economic purposes [347]. This is shown in Equation 14.

$$U = \frac{r}{C} \qquad (14)$$

With U as the utilisation of the system, and r as the rate of entry of product (which be the demand or load) and C as the system capacity.

Caution must be applied to avoid the intuitive "rule" that utilisation must be as high as possible. Designing systems by targeting absolute utilisation creates the risk of the system going unmanageably out of control [98]. Low utilisation means that a system is less capable of delivering a service, however as utilisation nears hundred percent (a practical impossibility, as constrained by Hopp & Spearman's laws of utilisation and capacity [40]), the system's ability to respond to demand falls drastically and queues are elongated uncontrollably [98]. Conscious of the harmful effects of very high utilisation, – the NHS has introduced an 85 % bed occupancy "rule" [192] as a means of keeping the system in control.

In the first paper of this suite of publications, we argued that demand in health care is structured as is shown in Equation 15.

$$D_t = D_v + D_e + D_f + D_n \tag{15}$$

With: D_t *: total demand,* D_v *: value demand,* D_e *: escalation demand,* D_f *: failure demand,* D_n *: false demand.*

The elements (or as we prefer "modalities") of demand shown in Equation 15 suggest a way to understand how the queue in a health system is constituted. We use Seddon's view that demand is an aggregation of value demand and failure demand [44]. However, we include modalities outside of his binary classification. We include false demand (which emerges from the healthy population) and escalation demand (the load placed on the system due to delayed treatment - for example, in the USA, it was found that one in eight cases became more severe due to delayed treatment [8]).

The capacity of a system is its ability to execute a function [347]. This capacity cannot be exceeded and only sustained for transient timeframes, as a multitude of limitations, called "detractors" reduce this capacity from the base capacity to the process capacity [40, 347].

$$C = C_b - D \tag{16}$$

With: C : Process capacity, C_b: Base capacity D: Detractors

Equation 16 [347] shows how detractors reduce the base capacity of a system to the true capacity, which becomes the target, which is in many cases the arbitrary product of average utilisation, one good shift and other factors [347]. Detractors may include failures, breakdowns, resource unavailability, start-up effects and others, which Bicheno calls "equipment losses" [38]. These elements can be managed to achieve zero losses [348]. Another category are so-called "dispensable-time-losses" [38] which include meaningless meetings, capturing and reporting on data that is never scrutinised, bureaucratic clutter, or what Graeber refers to as "BS jobs" [349][21].

[21] Graeber is less restrained in his use of the unabbreviated form, which the curious reader may find in the reference list.

5 - 3. Objectives and Approach

Recognising the impact that failure demand has in many industries, we set out to understand and define this phenomenon in health care. To do so, we aim to:

- build a logical framework that can be used to assess events to classify them as either "system failure" or not,
- use the developed framework to classify commonly occurring events in health care and identify them by modality,
- test the framework and the classifications with a panel of experts to validate the utility, primarily of the framework and secondarily of the classifications.

5 - 3.1. Method

Roy et al. say that concepts are incremental and build upon existing knowledge, ideas observations and their synthesis. [130]. To ensure that the framework presented in this paper is a credible tool, we surveyed a panel of experts to validate the usefulness of the framework and to ensure that it was a comprehensive treatment capable of delivering a valid conclusion of system failure.

To achieve our aims, we reviewed applicable global literature sources to identify the major drivers for patients leaving health services. We augmented these sources with several exploratory studies and general observations.

This study consisted of four major parts as shown in Figure 5-23:

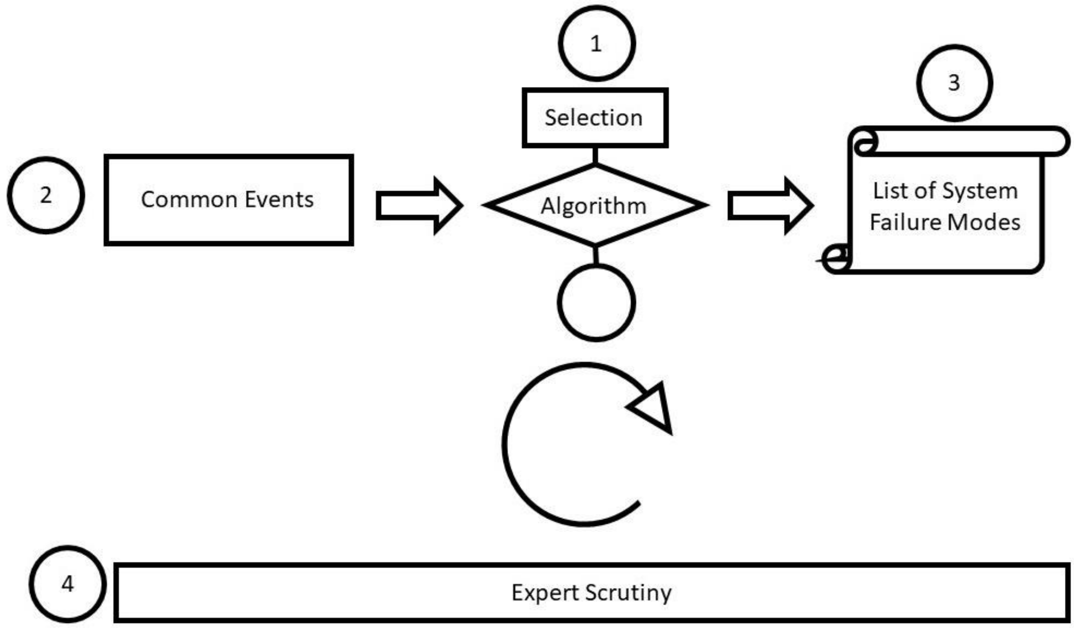


Figure 5-23: The phases of this study

Firstly a framework was constructed ① making use of literature sources, general observations and brainstorming [350]. The purpose of this framework was to serve as a logical test of a scenario to evaluate whether the event is system failure, or not.

Secondly, a selection of eighteen common events was compiled based on observations of failure in health systems from literature sources and our own experience ②. These events were presented to the framework developed above and in so doing tested the ability of the framework to correctly conclude whether an event was system failure, or not. Based on the application of each of the events to the model, a categorised list of events was presented ③.

Model Validation

Using Dijkstra et al.'s expert validation approach [350] a panel of experts was engaged to validate the framework as well as the findings of this study ④. Ten international experts in Lean health care were contacted by email to request their support in validating the model. Two of the experts ignored the request for support and one declined participation citing high COVID-19 workload. One of the experts re-shared the survey with an additional expert, thus bringing the total response rate to eight.

What was shared with the experts?

The experts were able to watch a recorded video which was hosted on a private YouTube channel. The video described all the major elements presented in this paper, with major sections devoted to "a recap of failure demand", "hierarchical models", "the development of the algorithm", "identifying the activities to be tested against the algorithm", "demonstration of the algorithm in use" and a "conclusion of the categorisation of events". The video was 27 minutes long, however watching speeds up to 1.5x remained feasible – and experts were encouraged to do so. A draft of this paper was also shared on request.

After watching the video, the experts accessed a Google-sheet which automated the collection of opinions. Three major sections were collection of key information about the experts, gathering opinion and inputs about the algorithm and finally conclusions about the findings of the study.

The outputs shown in this paper reflect two rounds of interactions with the panel of experts and represents the endpoint of considerable iteration to arrive at consensus.

Expert Credentials

All experts are prominent professionals in Lean health care. They have been sourced from academia (5) private consulting (2) and health system management (1). The experts' self-reported experience in Lean health care averaged to sixteen years. All experts have published extensively in the field, including high impact journal publications, keynote appearances and at least five published books between them. Four of the experts hold a PhD or are completing one. Two have a master's degree and two have undergraduate degrees. Six of the experts have been involved in Lean training, developing bespoke material for academic and private organisations. Six experts have led a major Lean project. Although the credentials of the experts are beyond question, it was interesting, that most experts were modest when assessing their own expertise, with the average, self-reported score of expertise being 3.75 on a 5-point scale Likert scale.

5 - 3.2. Expert assessment

The framework presented in this paper is the final product that has undergone expert validation. Modifications have been incorporated into Figure 5-24 and by extension Table 5-2: .

5 - 3.2.1. **Changes to the framework**

Initial language used the terminology “delinquent act”. Although this was meant to refer to the act, two experts were concerned about the pejorative implications of the word “delinquent” creating the impression of a “bad” patient, or a “bad” doctor or nurse. This was not the intention for two reasons – the one being, that the systems view does not hold the view of individual “fault” [15, 99] but rather tries to assess systemic questions, and secondly, the wording unwisely implied the focus to lie on “fault” rather than on “the act” [44]. As this was by no means the intention of this model, the term was dropped and replaced by “triggering act”.

All experts agreed that the framework was useful. In a separate question, they awarded it an average score of 3.75 to be able to be used for other, untested scenarios. This was lower than hoped and seemed to emerge from a concern that the framework needed revision to be more generalisable. One expert suggested that to be more generalisable, clinical language should be removed from the formulation. This would allow the framework to be useful not only in clinical settings, but perhaps also in other complex hierarchical systems such as government and the legal system. Clinical specific language was deleted and generalised.

5 - 3.2.1.1. The nodes

One expert suggested that an emphasis on capacity planning and resource allocation [351] should be split out as an additional node, and not be covered under the “catch-all system-design” node. Doing so, also brings in the strategic layer [174] into the algorithm as equally a contributor to system failure.

A previous category of "support services" was eliminated as the support services are inside the service system boundary and covered by existing nodes. Similarly, a previously existing node for incompetence was deleted and merged with the "delinquent act" – later the "triggering act" – due to there being no meaningful distinction between them.

Two of the more subjective nodes were enriched by creating smaller sub-frameworks, these were the "triggering act" making use of the traditional failure demand definition [44]. A further framework was added to assess whether resolution was achieved [44, 113, 122]. This was necessary, to adjudicate the difference between what Lillrank et al. call "care" and "cure" [241] . This should ensure that returning for care should not be interpreted as system failure.

5 - 3.2.1.2. The addition of scenarios for testing

Only one expert identified a common scenario that should be added to the assessment of the model, i.e., the patient leaving without being seen. This was included in Section 5 - 5.3.5.

5 - 4. The System Failure Framework

The framework is shown in Figure 5-24. The main framework is shown between the two subordinate frameworks indicated by ① and ②. These smaller frameworks are used to assist in reducing the subjectivity on two nodes, as indicated. The "Triggering act" is derived from the pure definition for failure demand – not doing something, or doing something wrong [44]. The next node speaks to the complex idea of resolution. This merges the thinking of Womack & Jones [121] and Seddon [137] that the completion of a service is to be interpreted from the point of view of the customer, but equally, needs to be measured against a good practice standard and determining whether this is appropriate for the type of service required [241].

On the main framework, the next node assesses system design, on the operational and strategic levels [44, 174, 320], followed by a node that assesses poor sharing of information, which again could be systemic in nature. The final node assesses whether departures were due to capacity planning and resource allocation issues [351].

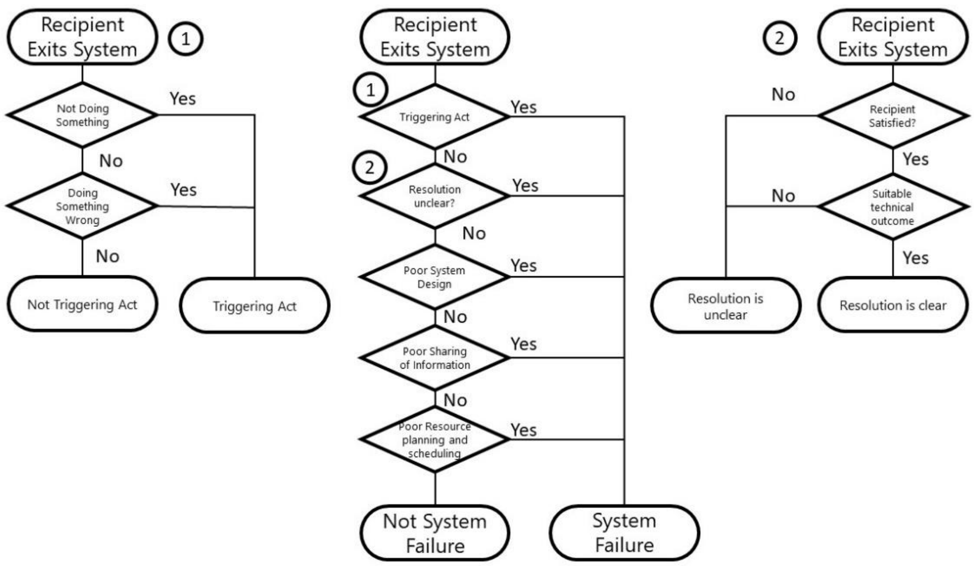


Figure 5-24: Framework for assessing system failure.

5 - 5. Testing the framework.

This section explores the modes of system failure in health care, which if accompanied by returning patients, can be classified as failure demand. System failure is the root cause of failure demand, and the manageable element in its reduction.

Sections 5 - 5.1 to 5 - 5.4 show the four key themes for which clarity is required to define system failure in health care. These deal with:

- the hierarchical nature of health systems,
- unsuccessful medicine,
- the errors made by patients,
- the overall operational environment for health provision.

5 - 5.1. Hierarchical structure of health care systems

A prominent difference between health- and other – systems are their complex, intentionally structured hierarchies. Underlying this structure are ideological, political and economic models in support of the health system [352]. Ranging from the day to day management to the strategic, system-wide design [174].

Most health systems are structured so that simpler care is provided at lower cost entities [353], such as in community information programs [238], at minor clinics, general practitioner practices, Health Management Organizations (HMOs) [220] and at outpatient departments. Higher skill and specialisation coincides with more costly and generally larger facilities such as specialised clinics and hospitals [345]. This protects highly specialised hospitals from overburden (Muri) [15] which is wasteful [90]. This structure also reduces the overall system cost, as primary interventions cost less.

To benefit from hierarchical levels, a referral system is used where patients arriving at a point of care of the incorrect level are referred to the correct facility. Patients who arrive at a level above their need, are referred downward, usually after some diagnostic and administrative work [354]. Similarly, patients, who arrive at a more primary point of care, are referred up through the health system until they arrive at the correct level of care, without unduly burdening more skilled, costly, and restricted higher diagnostic and administrative levels.

These hierarchies are probably necessary to provide health care, as up- and down – referral is a cost limiting mechanism by which patients arrive at the correct care level. It is plausible that patients will repeatedly exit and re-enter the health system, until they reach the correct level. Although this does unburden higher levels of care (and is favoured as the future direction of health care by Hopp & Lovejoy [211], we consider it as system failure, as upon their return, the health system must repeat the work already done on these patients.

We propose the relationship in Figure 5-25 to show system failures in complex systems, such as the health system. We do this, by modifying the previous model that represents the simple case as shown in Figure 5-22.

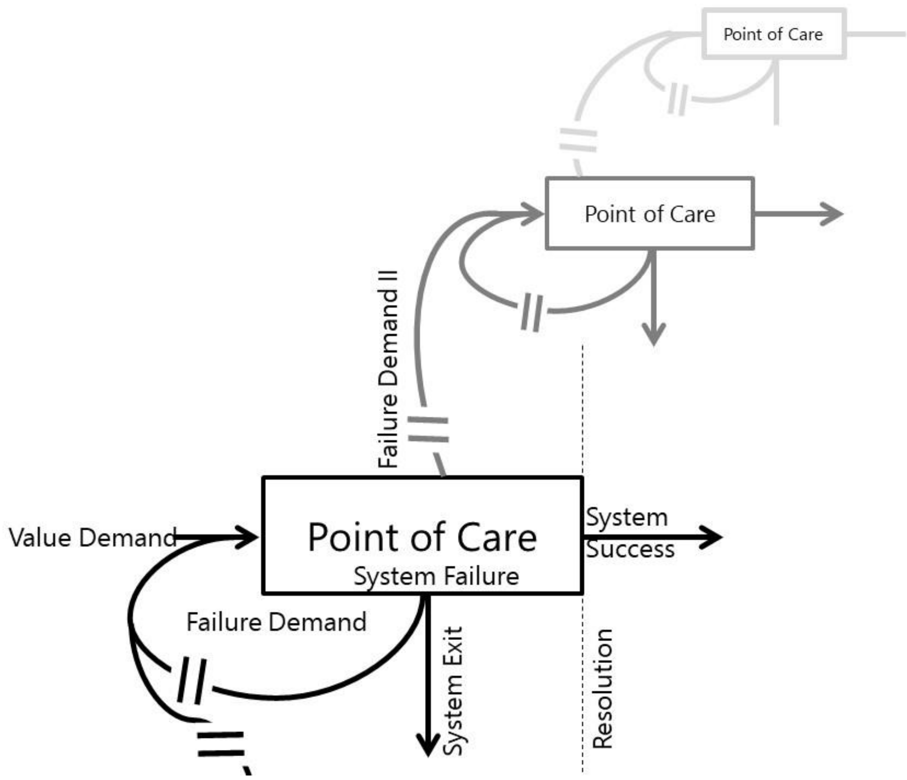


Figure 5-25 :The occurrence of system failure[22] and failure demand in hierarchical service systems, such as health care.

The failure demand model introduced in Figure 5-22 is repeated as the starting unit of the model above. The simple case assumes that system failures can only exit the system or return to the same point of care. The expanded model shows that system failures can cause failure demand at points of care other than the ones that created the system failure. Because the nature of this type of system failure is considerably different to the traditional system failure – failure demand relationships, we introduce the notation failure demand II.

[22] For ease of legibility, we only indicate failure demand II moving from a lower level point of care to a higher one, but it must be assumed that it equally moves down the hierarchy. The reader may also assume further levels of care above and below the presented model.

Failure demand II emphasises that load should be seen with a systemic lens [96] and that just because work has been moved from one point to another may not benefit the system as a whole.

Our expert-panel supported this concept, with only one outlier, citing practicality rather than correctness as their major concern. The expert assessment of this idea delivered an average score of 4 out of a possible 5.

Conclusion: System failure - system design -> repeated consumption.

5 - 5.2. Unsuccessful[23] medicine

Medicine can be unsuccessful for many reasons. Some forms are system failures, whilst others are not. This section describes different forms of unsuccessful medicine from literature and personal experience and provides arguments for categorisation. We discuss chronic care, experimental medicine, trial and error medicine, palliative care, and medical mistakes.

5 - 5.2.1. **Chronic care**

Chronic care is the care of diseases that "… are not passed from person to person. They are of long duration and generally slow progression. The four main types of non-communicable diseases are cardiovascular diseases (like heart attacks and stroke), cancers, chronic respiratory diseases (such as

[23] We introduce the concept of "unsuccessful" medicine with caution. As this section will show, the success of a medical intervention is not always defined by cure, as in the case of chronic or palliative care. We cannot however head this section as "non-curative" medicine, as that implies the intention not to cure, which belies categories such as experimental or trial and error interventions, which may not cure, yet strive to.

chronic obstructed pulmonary disease and asthma) and diabetes.” [355] [356, p. 2]

The nature of chronic care is that patients repeat on the system and as such, resolution in the simple sense has not been achieved. We propose that chronic care is not a system failure because the purpose of such care is not striving for a cure, but rather prioritising management of an ongoing condition [241]. Resolution in this case should be defined as the administration of an appropriate disease - management or diagnostic event - patient-exits are not system failures, but management events seen to resolution, that are programmed to return for further management.

Although chronic care is not an example of system failure, we propose that it should strive towards lower contact frequency, subject to the proviso that the health outcomes are not altered [357].

Conclusion: Not system failure - resolution is defined by “care”, not “cure”.

5 - 5.2.2. **Experimental medicine**

Medicine errs towards familiar approaches [346] yet treatment ranges from conservative interventions to invasive, and often unnecessary, but costly treatments [358]. Non-conservative treatment is sometimes inappropriate and even morally questionable, however, at other times, it is the only response to an unfamiliar medical condition.

When confronted with unseen problems and unfamiliar cases, clinicians may need to innovate in their treatment approaches. In Bangladesh, where innovation is encouraged, health indicators have the best trajectory in South Asia [359].

Nevertheless, the innovating clinician must choose approaches that fit the condition. It is unreasonable to deem experimental medicine as system failure, unless it is done contrary to good clinical practice. We raise the caveat that the experimental approaches should be limited to unfamiliar conditions and even then, benchmarked against best clinical practices.

Conclusion: Not system failure – scientific limitations may limit the ability of medicine to cure.

5 - 5.2.3. **Trial and error medicine**

We view trial and error medicine as a nuanced type of experimental medicine. This relates specifically to clinicians refining an intervention through iterative methods. An example is depression medication dosage. In general, patients will receive anti-depressant medication, and the dosage will be modified until a level is reached at which a clinical response is achieved [360]. This trial-and-error approach is a common, good clinical practice, and in our view not a system failure.

We add the caveat, that we view trial and error approaches as necessary but non-value adding activities [63], and the system must strive for greater knowledge and data to reduce the amount of trial and error required in medicine.

Conclusion: Not system failure – scientific limitations may limit the ability of medicine to cure.

5 - 5.2.4. **Palliative care**

At times, patients have no remaining prospect for improved health. Nevertheless they repeat on a health care system for palliative care, the maintenance of the best possible standard of living, which centres on comfort and dignity with no long term survival expectations [361].

Conclusion: Not System failure - resolution is defined by "care", not "cure".

5 - 5.2.5. **Medical errors**

Some deaths, are avoidable. Kohn Et al. [345] speak about the burden from medical errors in the USA; they mention two studies that show that around 3 % of patient interactions contain what they refer to as adverse events. Between 44 000 and 98 000 patients die per year in US hospitals due to medical errors. Although Hayward & Hofer [362] argue that the numbers of deaths are exaggerated, they do not dispute that medical errors are significantly problematic. Toussaint & Gerard [197] claim American clinicians make up to 15 million medical errors annually which range from incorrect drugs or dosages[24], incorrect site surgeries or infection. In the UK, the government reported that a million patients are "put in hospitals" annually due to medical errors [197, 363].

Attention to reducing medical errors in a systematic way can have dramatic results, as was famously shown at Allegheny General hospital, where the prevalence of central line infections was virtually eliminated [364] by using the principles of the Toyota Production System [15]. Similar improvements

[24] Drug errors can have two origins. Incorrect scripts can be written and the incorrect drugs can be given. In both cases, the drug itself or the dosage may be wrong [363]. Patients incorrectly taking medicines are separately treated in this text.

were made at Virginia Mason hospital, simply by raising awareness of errors that medical professionals were making unwittingly [157].

Even though they are seen as important, and are often deadly [365] incorrect diagnoses are generally not included in the definitions of medical errors in the literature, however we consider this false, ignoring root causes over symptoms. We include diagnosis mistakes, and therefore conclude that errors are even higher than stated.

The majority of failures occur as a result of poorly designed systems [239]. This emphasises the value of designing systems intentionally compared to poorly, or un-designed systems which allow (or even causes) errors. Although mistakes are "human" [345], Deming's 94-6 principle [239] observes that the smaller proportion of mistakes (6 % of mistakes) is due to human incompetence, negligence or malice, whilst the bulk of errors (94 % are accounted for by systemic issues).

The reasoning above tries to show that medical mistakes are system failures and that there is evidence that strengthened systems leads to sustainably fewer errors and less harm. Thus, patients fortunate enough to return after having experienced a medical error, should be counted as failure demand.

Conclusion: System failure – system design does not minimise errors.

5 - 5.3. Patient errors

Patients are vital actors and stakeholders in a health system. Their actions are as much due for scrutiny as are those of clinicians. Although the patient is not included in a strict interpretation of Seddon's definition, we view this as a logical extension. A large portion of system failure is due to patient error of some sort. This section will explore five typical patient errors, including patients arriving at the wrong site, or at the wrong time, patients who disobey pre-treatment instructions, or those who arrive with the incorrect paperwork.

5 - 5.3.1. **Patients who arrive at the incorrect point of care.**

Some health systems are structured into health districts, (e.g. South Africa and Spain), where patients have to be seen by the facility that serves the district in which they live [354, 366]. This means that a patient's address determines their health facility. When patients go to an "incorrect" point of care, they are eventually turned away. Patients might also arrive at the incorrect level [25] of care, however this is generally a clinical mode of system failure and is separately addressed under Section 5 - 5.1.

Conclusion: System Failure - system design and poor communication.

5 - 5.3.2. **Patients who arrive for care at the incorrect time**

Many health systems provide regularly scheduled care for certain conditions. For example regular Aids clinics in Entebbe, Uganda [367], regular Diabetes clinics for military veterans in Los Angeles, USA [368] and others. In general, these clinics are scheduled for particular days and patients who arrive on a different day do not receive the service.

Similarly, patients in scheduled care, who arrive at a time other than their appointment time will often not be seen. In the Spanish system, some

[25] As opposed to point of care, which has a geographic element.

primary emergency departments only operate during business hours, requiring patients to move to general hospital accident and emergency departments after hours [354]. We have frequently observed an end-of-day migration of many dozens of un-served patients from the outpatient - to the emergency - department.

Conclusion: System failure - system design and poor communication

5 - 5.3.3. **Patients who disobey pre-treatment instructions**

Many treatments require patient behaviours in preparation for treatment. For example, pre-surgical patients are required to “starve” prior to surgery [369]. In a pilot study in a surgical theatre, we found that 2 % of procedures are cancelled because patients had eaten inside the “starvation period” [370]. Patients must return on a future date for the same treatment. Similarly, some laboratory tests require a certain diet leading up to the actual test [371]. Non-conformance leads to repetition or incorrect results.

Conclusion: System failure - the system designed allows non-adherence and proper practice is inadequately communicated. Occasionally, there is a wilful triggering act.

5 - 5.3.4. **Patients who arrive with incorrect paperwork**

Patients are seldom seen without suitable forms of national identification. This is particularly true in countries that have essentially free (from the point of view of the patient) health care. Equally, patients subject to failure demand II, generally need to bring referral documents. These letters serve as the primary communication across levels in a health system [372] without which patients will often not be seen.

Conclusion: System failure – system design.

5 - 5.3.5. **Patients who take medication contrary to instruction**

The US – Surgeon General reported that 75 % of Americans have trouble taking their medications as directed [373, 374]. This is worrisome, especially when read together with Di Matteo et al.'s finding that clinical outcomes are three times worse in patients that have poor adherence [375]. These poor outcomes can include common adverse clinical side effects [376] or worse. McCarthy found that as many as 125 000 Americans die annually due to improperly taken medication [377], though Meredith emphasises that the true scale of the problem is unknown [159]. A variety of reasons exists for poor adherence [378], which includes polypharmacy (taking more than 5 medications daily (polypharmacy [379]) forgetfulness, unclear clinical instructions and high cost [380]. Non-adherence is particularly common among people who live alone [376] or are elderly [381].

Conclusion: System failure – instructions not adequately communicated, or lack of remedies to compensate for patient-driven non-adherence.

5 - 5.3.6. **Conclusion on patient errors**

We propose that all presented instances of patient errors are a form of system failure. According to Deming [239] a system should be designed in such a way that people cannot make mistakes. Therefore, although the error is that of the patient, the fault lies with the health system that did not adequately inform patients of their duties, obligations, and functions.

The patient errors shown here are a consequence of patients being equipped with inadequate information or inadequate understanding of available information. A system wide intervention to inform patients, familiarise them with the operational modes of the health system and encouraging compliance is a systemic intervention that could reduce failure demand. Simultaneously the health system should be reengineered to make it more difficult to make errors.

5 - 5.4. Operational environment

Medicine is in many ways idiosyncratic. This section evaluates these as potential systemic drivers for failure. We explore six common elements, namely: financing, queues, supply-chain, support services, staffing and infrastructure.

5 - 5.4.1. Patients who have insufficient financial means for treatment

Article 25 of the Universal Declaration for Human Rights [382] states the right to health and health care is inviolable. Backman Et al. express that the right to health is not only “good management, justice or humanitarianism” but indeed, that providing such care, is an “obligation under human rights law” [383, p. 2047]. Nevertheless, only few nations can provide fair access to health care despite more than sixty years of this principle being universally accepted[26].

[26] Regretfully, of the thirty articles of the Universal Declaration of Human Rights, not a single right among them has been adopted universally.

Fairness is one of the three pillars of the World Health Organization for health systems [353]. The notion of fairness includes access. Economic exclusion is a form of an unfair health system.

The cost of medical treatment is of global concern. In many OECD states, the objective of the health system, is to base care on need and not on means [384]. This study found that generally, European states are more able to provide care on this basis, whilst notably it was found that in the USA, health-seeking behaviour was significantly biased towards the wealthy.

In the USA 29% (Sarnak reports 33% [7]) of patients take medications incorrectly, due to the cost of consuming the medication at the correct rate [8], whilst 12 % of Americans cannot afford their medical bills [8]. Indeed, 64 % of Americans report "the fear of unexpected medical expenses" as their greatest financial worry, ahead of mobility, heat, utilities or having somewhere to stay [8]. As a result half of Americans report being discouraged from seeking medical care [8] leading to self-medication and conditions being ignored [218].

The introduction of the Affordable care Act in the USA [10], even in its naming, tried to address high costs of care. Whether or not it has achieved this objective is unfortunately mired in political discourse (see e.g. [385]) however, the intention of reducing health care cost is noted as a priority.

Compared to developed countries, developing countries face an even greater burden from insufficient finance leading to even more pronounced exclusion from health services. For example, the Chinese health system is designed with the intention that the patient pays for services, although a rapidly emerging health insurance industry exhibits complexity beyond the scope of this paper [386].

Hsiao [387] reported that the Chinese system resulted in the economic exclusion from care of poorer people. Hall Et al. [388] showed how diabetes treatment in Sub Saharan Africa impoverishes families, with Sudanese families spending up to two thirds of their income on caring for a diabetic child. Leive & Xu [389] show how in fifteen African states generally between 30% and 40 % (70 % in Burkina Faso) of people need to borrow money or sell property to cover health care expenses at some time. In Burkina Faso it was found that up to 15% of households suffer catastrophic health expenditure, and this for relatively low levels of care [390].

Perhaps no single measure as completely reflects the failure of the entire system, from its intent to its functionality, as the exclusion of those, who need care, for financial reasons. Nevertheless, this is a daily reality in many countries globally, leading to failure demand, particularly, because those so excluded could later return with exacerbated conditions. This places an elevated demand on the health system.

Conclusion: System failure in the specific sense and more broadly, as it refers to the overall purpose of health care in the first place.

5 - 5.4.2. **Queues.**

The management of queues in health care systems is a significant research field. Many studies have attempted to shorten these, either through improved efficiency or by planning capacity better [156]. One reason to focus on queue length, is that patients may choose to balk or leave if the queue for care is too long (leaving without being seen). In a simulated study at a real site in the USA, the losses due to balking amounted to over US$ 680 000 per month [391]. Methodologically, this is a transferable figure, suggesting that significant losses probably hold true in many similar environments.

Bottlenecks in system designs further accumulate load and reflect poor approaches to resource-management.

Strategically, Allder et al. [392] identify demand-lag strategies – where queues exist, but remain stable and are in-effect a buffer against system variation. Strategically [174], queues are only problematic if they continue growing, which would indicate a systemic under-capacity. Tactically however, balking due to predictable queue length is problematic [44] and steps should be taken to improve this.

The question of interest for this paper is whether long queues and consequent balking is system failure. The reason for long queues is definitive in answering this question. Queues are caused either by poor planning or by variation in load:

Conclusion: Both. Seddon speaks about predictability [44]. If the queue length is predictable, then the cause is systemic, then the problem is system failure. Unpredictable failure however (like a bus accident, or a stadium stampede) is not system failure.

5 - 5.4.3. **Supply-chains and inventory management**

Managing inventories of drugs, consumables, and other health care resources is of considerable importance.

In Blantyre, Malawi, it was found that a major reason for anti-retroviral non-adherence, was frequent stock-outs in the pharmacy [393]. The returning patient is not only failure demand, but also quite possibly immunocompromised, potentially escalating the severity of the illness.

In a pilot study in a surgical ward, we found that over 15 % of surgeries were cancelled due to clean linen not being available. The absence of competent supply-chains for necessary items led to the delay of treatment and the escalation of the severity of the condition and we view it as system failure.

Conclusion: System failure - poor system design and triggering act.

5 - 5.4.4. **Staff unavailability**

In another study, we found that the late arrival or the non-availability of nurses, doctors and anaesthetists accounted for roughly 30 % of surgical delays [370]. We found that in general, surgical days started over 90 minutes after their scheduled start [394] delaying the whole schedule and often leading to cancelled procedures. We have observed similar issues in General Practitioner (GP) and other practices, where the absence of a medical professional leads to system failure.

Conclusion: System failure - triggering act.

5 - 5.4.5. **Delays from support or diagnostic services.**

Delayed laboratory-results increased waiting time, and reduced overall system efficiency [395]. A study found that introducing laboratories into an emergency department, improved the productivity of the unit by the same quantum as hiring an additional nurse and clinician [396]. Similarly, improved laboratory turnaround times reduces overall waiting times, improves health outcomes, as clinicians can make evidence-based decisions.

Long waiting times for laboratory results are therefore doubly system failures, as waiting times are increased and clinicians may make mistakes due to occasional time-pressured interventions, when choosing to proceed with treatment before receiving delayed lab results.

Conclusion – System failure – poor system design.

5 - 5.4.6. **Lack of infrastructure**

Many health providers can only perform certain treatments if the patient can foreseeably be admitted to a hospital bed either prior to, or after, treatment. We found that 23 % of surgical cancellations were due to high- or intensive-care bed shortages required for post-operative admission [370]. These patients had often been admitted to general wards and starved in preparation for elective surgery, prior to cancellation.

Conclusion: System failure –system design and badly planned capacity and resource allocation.

5 - 6. System Failure in Health Care

5 - 6.1. Summary of system failure events

In the preceding sections, we have differentiated between events that amount to system failure, and those that do not. Table 5-2: summarises the reasoning from these sections. The reader may follow the section headers in parenthesis for a reminder of the argument in each case.

Recalling Equation 16, non-system-failure events represent detractors as much as they do value demand. In both instances, managing these cases and limiting their occurrence is possible and necessary, however they do not represent a failure of the system, and their returning demand would be considered value demand.

Broadly, a well-designed system is one which in our view has not failed. For example, Hopp & Lovejoy show a variety of benchmarks for health system design [211]. Based on these, bed numbers and system capacity are mandated. Should the demand on a well-designed system be lumpy or erratic [397] beyond the best judgement for its design – this is excusable as a capacity concern and its classification as system failure would be unreasonable.

Table 5-2: Framework defining system failure in health care.

System failure	Not System failure
Moving up or down the hierarchy (5 - 5.1)	Unsuccessful Medicine (5 - 5.2)
Medical errors[27] (5 - 5.2.5)	Chronic Care (5 - 5.2.1)
Patient errors (5 - 5.3)	Experimental medicine (5 - 5.2.2)
Wrong Site (5 - 5.3.1)	Trial and error medicine (5 - 5.2.3)
Wrong time or date (5 - 5.3.2)	Palliative care (0)
Disobeyed pre-treatment instructions (5 - 5.3.3)	Predicable Long Queues (5 - 5.4.2)
Incorrect paperwork (5 - 5.3.4)	
Operational Environment (5 - 5.4)	
Insufficient finances (5 - 5.4.1)	
Patients who take medication contrary to instruction (5 - 5.3.5)	
Unpredictable Queues (5 - 5.4.2)	
Supply-chain and Inventory Management (0)	
Staff Unavailability (5 - 5.4.4)	
Support and diagnostic service delays (0)	
Lack of infrastructure (5 - 5.4.6)	

5 - 7. Implications

5 - 7.1. Theoretical contributions

We expand failure demand to introduce system failure to better explain failure demand, and then present researchers and practitioners with a model for identifying and measuring system failure, and therefore by extension more holistically, failure demand, in health care systems.

We emphasise the causal relationship between system failures as the root cause and failure demand as the symptom. In our model, derived from

[27] Including incorrect diagnoses, drug errors, which includes wrong prescriptions and prescriptions incorrectly filled, and in both cases; the incorrect drug or inappropriate dosages. This category further includes wrong-site surgery and other medical errors.

literature and our experience, events that are system failure are identified, which upon a patient's return, becomes failure demand.

The framework comments on the hierarchical nature of system failure in health care, and so introduces the idea of failure demand II, the mode of failure demand that crosses hierarchical bridges. This enables an overall systems-view of health care facilities including interconnectedness among facilities, which has implications for regional and national policy [174].

On a philosophical level one expert asked the question "whether any distance from the ideal should be seen as system failure". We propose that the answer to this is "yes", and that this logic must be applied in the same way that non-value adding activities must be strictly identified [63], even if they are necessary or unavoidable. Our approach in categorising events was inspired by Lean best practice, which urges one to err on the side of severity when considering whether an activity is wasteful or not [63]. Defining an activity as value adding leads to its long-term classification, and remedies to improve are not sought. Following this reasoning, we tend to categorise ambiguous cases as system failure, rather than protecting such behaviours against future scrutiny.

5 - 7.2. Managerial implications

Failure demand has been found to range between 40 % and 80 % across a variety of industries. It would be interesting to establish what the impact of failure demand is in the health care industry. If the incidence of failure demand is high, that would provide an interesting insight into how much demand is avoidable.

Systematic interventions can reduce demand, meaning that capitalisation and staffing can be reduced, or more usefully, service levels could improve at no additional cost. This, in the face of considerable cost pressures on health systems globally, is desirable.

Given that reduced failure demand could dramatically impact demand patterns, we advocate that health system strengthening policy incorporates the reduction of system failures and considers failure demand as an important design consideration for strategic system design [174].

5 - 7.3. Societal impact

From common observations, we have identified four cases that occur post system failure. The first case is the classic case of failure demand, patients who have experienced system failure return for repeat service in pursuit of resolution. The second case represents those patients who do not return but get better by themselves, so called "self-limiting conditions" [307] The third case represents patients who do not return to care, and do not get better, but also do not get worse [390], What we refer to as the "suffering in silence" population. The last group of patients do not return for medical care and their conditions worsen and may be as severe as fatalities.

In the first case, the burden, which includes all costs of system failure is carried by the funder of medical care. In the other three cases, the full burden of system failure is carried by society.

To understand the reasons why patients transfer the burden to society, we propose that most patients who do not return, do so because they are poorly informed, and do not realise that their condition can be improved. If they believe that their condition can be improved, they have often come to mistrust the capabilities of health systems or have become discouraged in seeking care. Further considerations include the time they will have to take off work to seek care, and the financial means they require to pay for it.

5 - 7.4. Recommendation for further study

This work should be expanded in field trials exploring the areas of concern.

The impact of failure demand can now be measured in health care settings, this should be done.

Events that are not system failure for example the "care-spectrum" should be examined in further depth to find opportunities to create systems that unburden the health system even when "cure" is not the aim.

5 - 7.5. Conclusion

This paper contributes a clarified model for understanding system failure in complex hierarchical systems, such as health care. This model provides several opportunities for future work. Researchers can use the model to identify and measure failure demand in selected health care settings whilst policy makers can incorporate thinking around failure demand for health system planning and design.

5 - 8. References

7. Sarnak, D.O., Squires, D., Kuzmak, G., and Bishop, S., *Paying for Prescription Drugs around the World: Why Is the Us an Outlier.* Commonwealth Fund Issue Briefs, 2017. 2017, p. 1-14.
8. Dijulio, B., Kirzinger, A., Wu, B., and Brodie, M., *Data Note: Americans' Challenges with Health Care Costs.* Kaiser Family Foundation, 2017.
10. Affordable Care Act, *Public Law 111-148*, in *Title IV, x4207, USC HR*. 2010, United States of America. p. 2010.
13. Graban, M., *Lean Hospitals*. 2008, CRC Press, Boca Raton, London, New York
15. Liker, J.K., *The Toyota Way*. 2004, McGraw Hill, New York
38. Bicheno, J. *Towards Reducing Queues: Muri, Mura, Muda*. in *European Lean Educator Conference (ELEC2018)*. 2018. Braga, Portugal.
40. Hopp, W.J. and Spearman, M.L., *Factory Physics*. 2011, Waveland Press, Long Grove, Illinois
44. Seddon, J., *Freedom from Command and Control: Rethinking Management for Lean Service*. 2005, CRC press, Boca Raton, Florida
47. Little, J.D., *A Proof for the Queuing Formula.* Operations Research, 1961. 9(3), p. 383-387.
63. Bicheno, J. and Holweg, M., *The Lean Toolbox: A Handbook for Lean Transformations*. 5th Edition ed. 2016, PICSIE Books, Buckingham, UK
90. Womack, J.P. and Jones, D.T., *Lean Thinking: Banish the Waste and Create in Your Corporation*. 1996, Simon and Schuster, London
96. Checkland, P., *Systems Thinking.* Rethinking Management Information Systems, 1999, p. 45-56.
98. Bicheno, J., *The Service Systems Toolbox: Integrating Lean Thinking, Systems Thinking and Design Thinking*. 2012, Picsie Books, Buckingham, UK
99. Senge, P.M., *The Fifth Discipline.* Measuring Business Excellence, 1997. 1(3), p. 46-51.
112. Sasser, W.E., *Match Supply and Demand in Service Industries.* Harvard Business Review, 1976. 54(6), p. 133-140.
113. Levitt, T., *Production-Line Approach to Service.* Harvard Business Review, 1972. 50(5), p. 41-52.

121. Jones, D.T. and Womack, J.P., *Lean Consumption.* Harvard Business Review, 2005. 83, p. 58-68.
122. Womack, J., *Lean Consumption.* Manufacturing Engineer, 2005. 84(4), p. 5-5.
130. Roy, R.B., Lillrank, P., Sreekanth, V., and Torkki, P., *The Conceptual Tools*, in *Designing Service Machines*. 2019, Springer. p. 5-19.
131. Roy, R.K., *A Primer on the Taguchi Method*. 2010, Society of Manufacturing Engineers, Ontario, Canada
134. Seddon, J. and Brand, C., *Debate: Systems Thinking and Public Sector Performance.* Public Money & Management, 2008. 28(1), p. 7-9.
136. Seddon, J., *I Want You to Cheat!: The Unreasonable Guide to Service and Quality in Organisation*. 1 ed. 1992, Vanguard Press, Buckingham, UK
137. Seddon, J., O'donovan, B., and Zokaei, K., *Rethinking Lean Service*, in *Service Design and Delivery*. 2011, Springer. p. 41-60.
141. Radnor, Z.J., Holweg, M., and Waring, J., *Lean in Healthcare: The Unfilled Promise?* Social Science & Medicine, 2012. 74(3), p. 364-371.
152. Hartmann, D. and Mandavha, R. *Lean Healthcare, a Casualty of Inefficiency*. in *SAIIE Conference*. 2010. Pretoria, Southern African Institute for Industrial Engineering.
156. Silvester, K., Lendon, R., Bevan, H., Steyn, R., and Walley, P., *Reducing Waiting Times in the Nhs: Is Lack of Capacity the Problem?* Clinician in Management, 2004. 12(3), p. 105-109.
157. Kenney, C., *Transforming Health Care: Virginia Mason Medical Center's Pursuit of the Perfect Patient Experience*. 2012, CRC Press, Boca Raton, Florida
159. Meredith, J.O., Grove, A.L., Walley, P., Young, F., and Macintyre, M.B., *Are We Operating Effectively? A Lean Analysis of Operating Theatre Changeovers.* Operations Management Research, 2011. 4(3-4), p. 89.
174. Hines, P., Holweg, M., and Rich, N., *Learning to Evolve: A Review of Contemporary Lean Thinking.* International Journal of Operations & Production Management, 2004. 24(10), p. 994-1011.
185. Marr, B. and Neely, A., *Managing and Measuring for Value: The Case of Call Centre Performance*. 2004, Cranfield School of Management and Fujitsu, Cranfield, UK.

186. Walley, P., *Does the Public Sector Need a More Demand-Driven Approach to Capacity Management?* Production Planning & Control, 2013. 24(10-11), p. 877-890.
192. Allder, S., Walley, P., and Silvester, K., *Is Follow-up Capacity the Current Nhs Bottleneck?* Clinical Medicine, 2011. 11(1), p. 31.
194. Walley, P., Found, P., and Williams, S., *Failure Demand: A Concept Evaluation in Uk Primary Care.* International Journal of Health Care Quality Assurance, 2019. 32(1), p. 21-33.
197. Toussaint, J. and Gerard, R., *On the Mend: Revolutionizing Healthcare to Save Lives and Transform the Industry*. 2010, Lean Enterprise Institute, Cambridge, Massachusetts
198. Wagstaff, A., Van Doorslaer, E., and Paci, P., *On the Measurement of Horizontal Inequity in the Delivery of Health Care.* Journal of Health Economics, 1991. 10, p. 169-205.
207. Hartmann, D., *Improvement in Operating Theatre Efficiency through Better Measurement and Scheduling Msc. Research Report*. 2012, University of the Witwatersrand Press, Johannesburg.
211. Hopp, W.J. and Lovejoy, W.S., *Hospital Operations: Principles of High Efficiency Health Care*. 2012, FT Press, Upper Saddle River, New Jersey
218. Osborn, R., Squires, D., Doty, M.M., Sarnak, D.O., and Schneider, E.C., *In New Survey of Eleven Countries, Us Adults Still Struggle with Access to and Affordability of Health Care.* Health Affairs, 2016. 35(12), p. 2327-2336.
220. Tajeu, G., *Health Maintenance Organization (Hmo)*, in *The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society*. 2014, John Wiley & Sons, Ltd, Hoboken, New Jersey.
238. United States Department of Health Human Services, *Healthy People 2000: National Health Promotion and Disease Prevention Objectives*, in *Healthy People 2000: National Health Promotion and Disease Prevention Objectives*. 1991, US Government Printing Office.
239. Deming, E., *Out of the Crisis. Quality, Productivity and Competitive Position. Cambridge, Ma: Center for Advanced Engineering Study*. 1986, Cambridge University Press.
241. Lillrank, P., Groop, P.J., and Malmström, T.J., *Demand and Supply–Based Operating Modes—a Framework for Analyzing Health Care Service Production.* Milbank Quarterly, 2010. 88(4), p. 595-615.

307. Moerman, D.E., *Meaning, Medicine, and the" Placebo Effect"*. Vol. 28. 2002, Cambridge University Press Cambridge
320. Seddon, J. and Caulkin, S., *Systems Thinking, Lean Production and Action Learning.* Action Learning: Research and Practice, 2007. 4(1), p. 9-24.
339. Levine, R.J., *Institutional Review Boards.* BMJ: British Medical Journal, 1989. 298(6683), p. 1268.
340. World Medical Association, *World Medical Association Declaration of Helsinki. Ethical Principles for Medical Research Involving Human Subjects.* Bulletin of the World Health Organization, 2001. 79(4), p. 373.
341. Wallenius, J.J., *Failure Demand at Telefonbanken, Seb*. 2011, Chalmers University of Technology, Göteborg, Sweden.
342. Piercy, N. and Rich, N., *High Quality and Low Cost: The Lean Service Centre.* European Journal of Marketing, 2009. 43(11/12), p. 1477-1497.
343. Pidgeon, C., *Systems Thinking and Public Sector Efficiencies*. 2010, Northern Ireland Assembly, Research and Library Service, Northern Ireland.
344. Plsek, P.E. and Greenhalgh, T., *The Challenge of Complexity in Health Care.* BMJ: British Medical Journal, 2001. 323(7313), p. 625.
345. Kohn, L.T., Corrigan, J.M., and Donaldson, M.S., *To Err Is Human: Building a Safer Health System*. Vol. 627. 2000, National Academies Press, Washington D.C.
346. Sox, H.C., *Medical Decision Making*. 1988, ACP Press, Sydney, Australia
347. Hopp, W.J., *Supply Chain Science*. 2011, Waveland Press, Long Grove, Illinois
348. Mccarthy, D. and Rich, N., *Lean Tpm: A Blueprint for Change*. 2015, Butterworth-Heinemann, Oxford, UK
349. Graeber, D., *On the Phenomenon of Bullshit Jobs: A Work Rant.* Strike Magazine, 2013. 15(3), p. 1-5.
350. Dijkstra, J., Galbraith, R., Hodges, B.D., Mcavoy, P.A., Mccrorie, P., Southgate, L.J., et al., *Expert Validation of Fit-for-Purpose Guidelines for Designing Programmes of Assessment.* BMC Medical Education, 2012. 12(1), p. 1-8.
351. Hans, E.W., Van Houdenhoven, M., and Hulshof, P.J., *A Framework for Healthcare Planning and Control*, in *Handbook of Healthcare System Scheduling*. 2012, Springer. p. 303-320.

352. Frenk, J., *Dimensions of Health System Reform.* Health Policy, 1994. 27(1), p. 19-34.
353. World Health Organization, *The World Health Report 2000: Health Systems: Improving Performance*. 2000, World Health Organization
354. Sempere-Selva, T., Peiró, S., Sendra-Pina, P., Martínez-Espín, C., and López-Aguilera, I., *Inappropriate Use of an Accident and Emergency Department: Magnitude, Associated Factors, and Reasons—an Approach with Explicit Criteria.* Annals of Emergency Medicine, 2001. 37(6), p. 568-579.
355. World Health Organization. *Noncommunicable Diseases*. Health Topics 2015 2017-06-26 [cited 2017 2017-06-26].
356. Bernell, S. and Howard, S.W., *Use Your Words Carefully: What Is a Chronic Disease?* Frontiers in Public Health, 2016. 4, p. 159.
357. Wagner, E.H., Grothaus, L.C., Sandhu, N., Galvin, M.S., Mcgregor, M., Artz, K., et al., *Chronic Care Clinics for Diabetes in Primary Care.* A system-wide Randomized Trial, 2001. 24(4), p. 695-700.
358. Wennberg, J.E., *Dealing with Medical Practice Variations: A Proposal for Action.* Health Affairs, 1984. 3(2), p. 6-32.
359. Das, P. and Horton, R., *Bangladesh: Innovating for Health.* The Lancet, 2013. 382(9906), p. 1681.
360. Paykel, E.S. and Priest, R.G., *Recognition and Management of Depression in General Practice: Consensus Statement.* BMJ : British Medical Journal, 1992. 305(6863), p. 1198-1202.
361. Jin, J., *Clinicians Examine Advances and Challenges in Improving Quality of End-of-Life Care in the Icu.* JAMA, 2013. 310(23), p. 2493-2495.
362. Hayward, R.A. and Hofer, T.P., *Estimating Hospital Deaths Due to Medical Errors: Preventability Is in the Eye of the Reviewer.* JAMA, 2001. 286(4), p. 415-420.
363. Smyth, C., *Million Patients a Year Put in Hospital by Medicine Mistakes*, in *The Sunday Times*. 2017, The Times.
364. Shannon, R.P., Frndak, D., Grunden, N., Lloyd, J.C., Herbert, C., Patel, B., et al., *Using Real-Time Problem Solving to Eliminate Central Line Infections.* The Joint Commission Journal on Quality and Patient Safety, 2006. 32(9), p. 479-487.
365. Kirch, W. and Schafii, C., *Misdiagnosis at a University Hospital in 4 Medical Eras Report on 400 Cases.* Medicine, 1996. 75(1), p. 29-40.
366. Van Rensburg, H., *Health and Health Care in South Africa*. 2004, Van Schaik Publishers, Cape Town

367. Watera, C., Todd, J., Muwonge, R., Whitworth, J., Nakiyingi-Miiro, J., Brink, A., et al., *Feasibility and Effectiveness of Cotrimoxazole Prophylaxis for Hiv-1-Infected Adults Attending an Hiv/Aids Clinic in Uganda.* JAIDS Journal of Acquired Immune Deficiency Syndromes, 2006. 42(3), p. 373-378.
368. Ho, M., Marger, M., Beart, J., Yip, I., and Shekelle, P., *Is the Quality of Diabetes Care Better in a Diabetes Clinic or in a General Medicine Clinic?* Diabetes Care, 1997. 20(4), p. 472-475.
369. Maclean, A. and Renwick, C., *Audit of Pre-Operative Starvation.* Anaesthesia, 1993. 48(2), p. 164-166.
370. Author Withheld, *Reference Withheld for Reviewing Purposes.* 2020.
371. Kale, V.P., Joshi, G.S., Gohil, P.B., and Jain, M.R., *Effect of Fasting Duration on Clinical Pathology Results in Wistar Rats.* Veterinary Clinical Pathology, 2009. 38(3), p. 361-366.
372. Grol, R., Rooijackers-Lemmers, N., Van Kaathoven, L., Wollersheim, H., and Mokkink, H., *Communication at the Interface: Do Better Referral Letters Produce Better Consultant Replies?* British Journal of General Practice, 2003. 53(488), p. 217-219.
373. Benjamin, R.M., *Medication Adherence: Helping Patients Take Their Medicines as Directed.* Public Health Reports, 2012. 127(1), p. 2-3.
374. Cutler, D.M. and Everett, W., *Thinking Outside the Pillbox—Medication Adherence as a Priority for Health Care Reform.* New England Journal of Medicine, 2010.
375. Dimatteo, M.R., Giordani, P.J., Lepper, H.S., and Croghan, T.W., *Patient Adherence and Medical Treatment Outcomes a Meta-Analysis.* Medical care, 2002, p. 794-811.
376. Ellenbecker, C.H., Frazier, S.C., and Verney, S., *Nurses' Observations and Experiences of Problems and Adverse Effects of Medication Management in Home Care.* Geriatric Nursing, 2004. 25(3), p. 164-170.
377. Mccarthy, R., *The Price You Pay for the Drug Not Taken.* Business and Health, 1998. 16(10), p. 27.
378. Burnier, M., *Medication Adherence and Persistence as the Cornerstone of Effective Antihypertensive Therapy.* American Journal of Hypertension, 2006. 19(11), p. 1190-6.
379. Bedell, S.E., Jabbour, S., Goldberg, R., Glaser, H., Gobble, S., Young-Xu, Y., et al., *Discrepancies in the Use of Medications: Their Extent*

and Predictors in an Outpatient Practice. Archives of Internal Medicine, 2000. 160(14), p. 2129-2134.
380. Fund Commonwealth, *International Health Policy Survey of Sicker Adults.* Funding for Germany was provided by the German Institute for Quality and Economic Efficiency in Health Care, 2005.
381. Mulhem, E., Lick, D., Varughese, J., Barton, E., Ripley, T., and Haveman, J., *Adherence to Medications after Hospital Discharge in the Elderly.* International Journal of Family Medicine, 2013. 2013.
382. Un General Assembly, *Universal Declaration of Human Rights.* UN General Assembly, 1948.
383. Backman, G., Hunt, P., Khosla, R., Jaramillo-Strouss, C., Fikre, B.M., Rumble, C., et al., *Health Systems and the Right to Health: An Assessment of 194 Countries.* The Lancet, 2008. 372(9655), p. 2047-2085.
384. Hurst, J., *Performance Measurement and Improvement in Oecd Health Systems: Overview of Issues and Challenges.* Measuring up: improving health system performance in OECD countries, 2002.
385. Kaplan, T. and Pear, R., *New York Edition with the Headline: Unity Is Elusive as G.O.P. Presses Health Overhaul*, in *New York Times*. 2017, Arthur Ochs Sulzberger Jr., New York City. p. 1.
386. Blumenthal, D. and Hsiao, W., *Privatization and Its Discontents—the Evolving Chinese Health Care System*. 2005, Mass Medical Soc.
387. Hsiao, W.C., *The Chinese Health Care System: Lessons for Other Nations.* Social Science & Medicine, 1995. 41(8), p. 1047-1055.
388. Hall, V., Thomsen, R.W., Henriksen, O., and Lohse, N., *Diabetes in Sub Saharan Africa 1999-2011: Epidemiology and Public Health Implications. A Systematic Review.* BMC Public Health, 2011. 11(1), p. 564.
389. Leive, A. and Xu, K., *Coping with out-of-Pocket Health Payments: Empirical Evidence from 15 African Countries.* Bulletin of the World Health Organization, 2008. 86, p. 849-856C.
390. Su, T.T., Kouyaté, B., and Flessa, S., *Catastrophic Household Expenditure for Health Care in a Low-Income Society: A Study from Nouna District, Burkina Faso.* Bulletin of the World Health Organization, 2006. 84, p. 21-27.
391. Broyles, J.R. *Estimating Business Loss to a Hospital Emergency Department from Patient Reneging by Queuing-Based Regression*. in *IIE Annual Conference. Proceedings*. 2007. Institute of Industrial and Systems Engineers (IISE).

392. Balderston, D., Gonzalez, M., and López, A.M., *Encyclopedia of Contemporary Latin American and Caribbean Cultures*. 2000, Routledge, Abingdon, UK
393. Van Oosterhout, J.J., Bodasing, N., Kumwenda, J.J., Nyirenda, C., Mallewa, J., Cleary, P.R., et al., *Evaluation of Antiretroviral Therapy Results in a Resource-Poor Setting in Blantyre, Malawi.* Tropical Medicine & International Health, 2005. 10(5), p. 464-470.
394. Hartmann, D. and Sunjka, B.P., *Private Theatre Utilisation in South Africa: A Case Study.* South African Medical Journal, 2013. 103(5), p. 285 - 287.
395. Hannan, E.L., Giglio, R.J., and Sadowski, R.S. *A Simulation Analysis of a Hospital Emergency Department*. in *Proceedings of the 7th conference on Winter simulation-Volume 1*. 1974. ACM.
396. Paul, S.A., Reddy, M.C., and Deflitch, C.J., *A Systematic Review of Simulation Studies Investigating Emergency Department Overcrowding.* SIMULATION, 2010. 86(8-9), p. 559-571.
397. Ghobbar, A.A. and Friend, C.H., *Sources of Intermittent Demand for Aircraft Spare Parts within Airline Operations.* Journal of Air Transport Management, 2002. 8(4), p. 221-231.

Chapter Six: **Measuring Failure Demand in A National Pharmaceutical Supply-Chain**

6 - 1. Abstract

Purpose

The exploration of failure demand in health care is limited despite it being a significant demand-portion in other industries. We test the prevalence of failure demand in a selected health care setting namely a national pharmaceutical supply-chain in a developing country.

Method

This exploratory study, consisting of two cases, investigates pharmaceutical wholesale- and distribution-networks by using custom-designed data mining software to interrogate transactional records from the existing ERP system. This produces standard performance metrics and reveals baseline values for the potential impact of failure demand on service delivery.

Findings

In the wholesale network, a failure demand level of 56 % was found whilst in the distribution network, this value was 29 %.

Practical and social implications

The problems experienced in service delivery in developing countries is exacerbated by failure demand. We propose that more efficient service delivery is possible by designing procurement policies to reduce failure demand. The reduction of load on the system can equal, and probably exceed, the amount of failure demand discovered.

Originality and value

Failure demand has not been described in pharmaceutical supply-chains and this paper tests the models that have previously been proposed. The paper also shows the impact of failure demand in health settings and as such provides an insight into the unnecessary burden placed on a particularly strained system and the need to consider demand in strategic health-system design.

Key Words: Failure demand; Health Care; System Failure, Pharmaceutical Supply-Chains, Demand Management, Quality, Health Policy.

Article Classification: Research paper.

6 - 2. Introduction

This paper uses a framework that we had previously designed to evaluate the impact of failure demand on a health care setting. In this case, due to convenience factors, the selected environment is a national pharmaceutical supply-chain network. In doing so, we present the design of an analytical tool to measure failure demand.

This topic explores a topical issue of service delivery, in a multi-billion dollar business, namely pharmaceutical distribution. Failure demand is a system-dynamic that exists in service systems, which elevates demand levels through a previous occasion of “not doing something, or doing something wrong” [44 Pg. 76]. Failure demand can compromise the “moment of truth” [44] and lead to reduced customer satisfaction but also overburden and unevenness in the service system. This in turn leads to elevated workloads, over-resourcing, and downstream behaviours such as over-procurement and hedging and hoarding. Considering failure demand reducing strategies as a part of strategic health system designs [174] could meaningfully impact the ability of health systems to meet the needs of society.

6 - 2.1. Objectives

Failure demand can account for as much as 80 % of demand in some industries [44], yet the literature on this topic in health care is sparse. Failure demand affects the customer experience of service, the system resources consumed, the workload on the system and, if well understood, the strategic system design. We aim to investigate this topic in a health setting by:

- Developing custom data-mining software to assess failure demand in supply networks, with the aim of being re-usable in varied settings for current and future research,
- To present a baseline value for the impact of failure demand in a selected branch of the health system,
- To identify root causes and drivers for failure demand,
- To propose remedies to reduce the incidence of failure demand in this setting.

6 - 3. Context

This study investigated the phenomenon of failure demand by analysing the efficiency of the national pharmaceutical supply-chain in an emerging country. The purpose of this study was to create an awareness of the phenomenon and to illustrate its systemic effect on service delivery and overall efficiency.

The supply-chain architecture (Figure 6-26, shown enlarged in Appendix 4 Page 338) consisted of a diversity of inbound suppliers, that includes international suppliers via air or sea and regional and domestic suppliers, mainly by road. These feed a single main warehousing hub, that in turn supplies twenty secondary hubs, each of which in turn supply a network of several hundred retail customers. These retail customers (sites) can be pharmacies at hospitals, clinics or independent retail pharmacies. The sites ultimately serve patients.

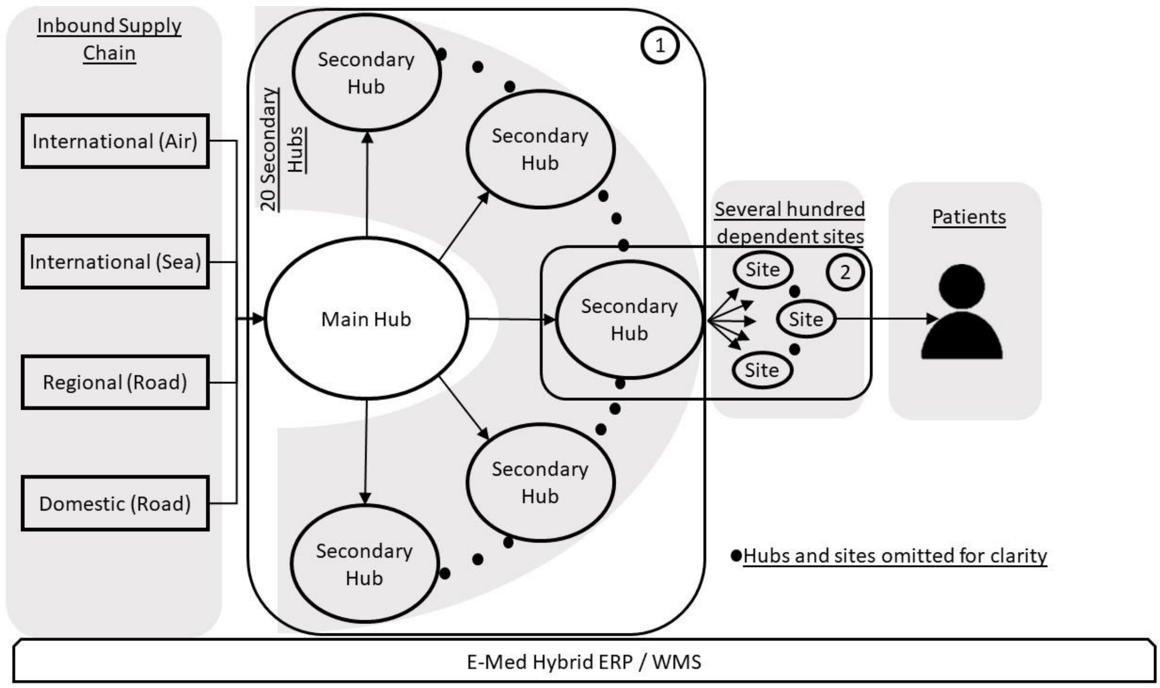


Figure 6-26: Supply-Chain Architecture

This study explores the supply-dynamics in the wholesale network between the main hub and the secondary hubs (indicated in Figure 6-26 by ①) and separately in the distribution network dynamics between a selected secondary hub and all dependent sites (or interchangeably, "health facilities", indicated by ②). The retail network, the relationship between sites and patients, is out of scope.

6 - 3.1. Processes

The system makes use of a hybrid Enterprise Resource Planning (ERP) – Warehouse Management (WM)-system, which we will refer to as "E-Med". E-Med is devolved from the main hub to sites, with orders and deliveries captured at all levels. The system performs numerous functions, including stock- and financial-control, order processing and administration. E-Med also manages manual warehouse operations. E-Med has no automated procurement capabilities, and no forecasting modules are in use, nor is the capability to manage transportation operations activated.

Figure 6-27 shows the processes that take place when an order arrives at the system. Most order-requisitions are manually delivered by customers, who remain at the warehouse, or nearby – as order fulfilment may take several days. The orders are captured in E-Med. Based on stock availability and some discretionary powers at the hubs, orders are modified to reflect the ability of the hub to supply to requirements. Where shortages occur for sites, letters of shortage are prepared which allow the customers to procure non-supplied goods on the open market should they wish to. Without such a letter, open-market purchases are illegal. Clients are not however compelled to purchase these products on the open market, they may attempt to order again from the hub later, or they may abandon the order. Goods that can be supplied, are picked from the warehouse, staged, checked, packed into the vehicle of the customer and dispatched.

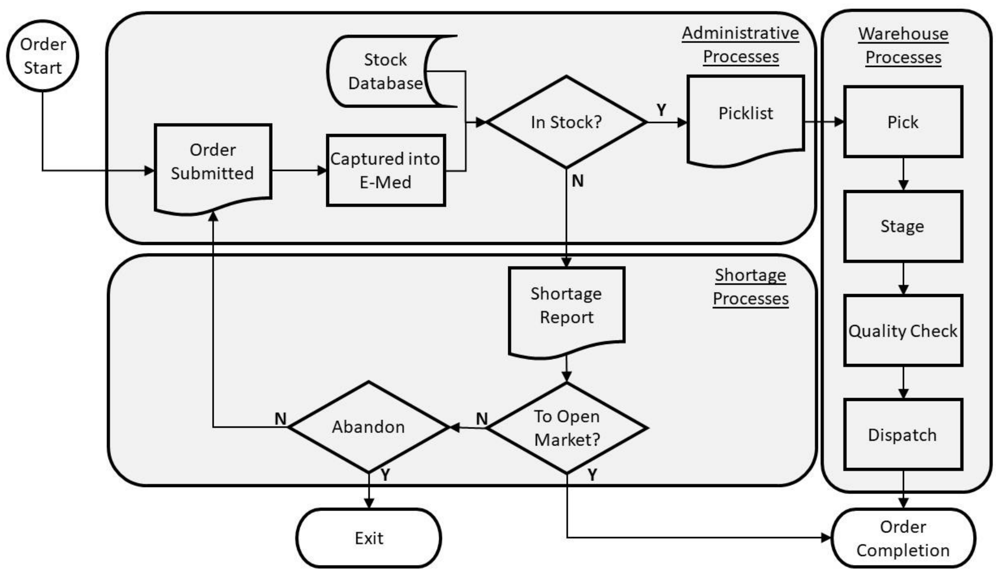


Figure 6-27: The process mapped from order arrival to system- completion or -exit

6 - 4. Literature

A strong link exists between the behaviours of people within systems and the measures employed to monitor and control them. Drucker says that improvement needs measurement, [398] whilst Goldratt speaks about the idea that people (or actors within a system) will behave in alignment with the measures used for evaluation [50]. The consequence of this is that (sometimes problematic) behaviours are driven by the evaluative framework in place. For example, in a system where reward is based on savings created by a manager – maintenance operations are often curtailed – to the long-term detriment of the system. Badly set targets can be achieved by gaming the numbers for compliance and bonuses [399]. This is characteristic of what Seddon calls a "command and control" culture [44], in which improvement is curtailed by performance metrics.

To emphasise the point, Muller expresses in 'The Tyranny of metrics' how measurement cannot be a substitute for judgement, but rather that measurement requires judgement [400]. Similarly, already in the sixties, Russell Ackoff cautioned against poorly designed metrics, titling his paper "Management misinformation systems" *[100]*. In the seventies, Goodhart's law *[401]* showed how metrics that become targets are inherently dangerous *[402]*. Strathern supports this, claiming that "audits can have a life of their own, threatening the life that it audits" *[402 Pg. 305]* . Thus, cautiously, this paper presents thinking around measures, without emphasising the suitability of them.

6 - 4.1. Traditional warehouse management

The Supply Chain Operations Reference (SCOR) model proposes several key metrics for assessing and managing the efficient operations of supply-chains, with a focus on warehousing tasks [403]. Broadly, these metrics have a strong quantitative base [404] and are structured into five key factors, namely : Reliability, Responsiveness, Agility, Cost and Asset Management.

This can be compared to the model of the Council of Supply-Chain Management [405] which proposes four phyla, namely Time, Quality, Cost and Uncategorised criteria. In both models, the similarity to Slack's model for competitive advantage is evident, with the categories being Quality, Speed, Cost, Flexibility and Reliability [406].

As can be seen in Table 6-3 (and reproduced for better legibility in Appendix 9 - 7), there is considerable alignment between the high levels of the assessment models, with the functional level aligning rather well, though classified differently by different authors, where similar content presented in different headings, these terms were underlined.

These metrics permit consistent trans-industry comparison against established benchmarks [407] which may lead to uniformity in practice [408]. They allow for efficient management and quality improvement. Although the relative importance of the metrics is highly context dependent, one study proposed that the most important metrics were Logistics Cost, On Time Delivery, Long Term Cost Trends, Customer Satisfaction and Stockouts [408]. We do however want to emphasise

Muller's view that for any metric, judgement in interpretation is required [400].

Table 6-3: Synthesis of performance management in supply-chain thinking. Modified from [403, 405, 406, 409-411]

Dimensions for competitive advantage [406]	SCOR Model [403]	Keeping score [405]	Metrics
Quality	-	Quality	Customer Satisfaction, Processing Accuracy, Order Fulfilment (on time, in full, accurate selection, damage free) Inventory Accuracy, Forecast Accuracy, Planning Accuracy
Cost	Cost	Cost	Inventory turns, sales outstanding, cost to serve, cash to cash cycle, total costs (cost of goods, transportation, carrying cost, MH costs, capacity costs, lost opportunity costs. Cost of goods sold,
Speed	Responsiveness	Time	On time delivery, order cycle time, cycle time variability, response time, forecasting cycle time, order lead times
Flexibility	Flexibility	-	Supply-chain response times, production flexibility.
Dependability	Reliability		Delivery Performance, Fill Rates, order fulfilment.
-	Asset Management Efficiency	-	Cash to cash cycle times, Stock Reach, Asset turns.
-	-	Uncategorised.	Minimum order Quantities, Change order Timing

As a strategic tool, however, these measures are not generally useful [412]. They are not directive in terms of future operations but rather prescriptive of a current state that should be managed tactically and operationally.

6 - 4.2. Good practices

Good Distribution Practices (GDP), Good Warehousing Practices (GWP) and Good Manufacturing Practices (GMP) exist in many industries, and the World Health Organization has proposed non-binding best-practice discipline-specific guidelines for the warehousing and transportation of pharmaceuticals [413]. These guidelines are designed to support continuous improvement. GDP focuses on seventeen areas for operational excellence and has a total of 179 distinct assessment points. These areas are generally qualitatively assessed and require an assessor who must have sufficient experience to be consistent. The assessor should also have independence, to minimise observers' paradox [294]. Generally, conformance to the criteria is assessed on a Likert scale and aggregated for an overall score.

6 - 4.3. System design

Warehousing and logistics are Necessary, 'Non-Value-Adding Activities' (NNVA [414]). That implies that if a way could be found to supply product to customers without warehouses or supply-chains, the customer would not complain. Therefore, warehousing and logistics networks must be designed for minimal cost and maximal efficiency. Warehouses are generally built to deal with current demand and a reasonable growth expectation over a certain period.

Understanding demand, allows the design of a logistics network of suitable size, with appropriate staffing and technology mixes for lowest cost, highest efficiency operations [409], which includes the correct sizing of physical infrastructure, computer systems and the workforce [415]. Working in demand-uncertain domains is always complex, subject to numerous internal and external stimuli. Capacity planning may require sophisticated approaches to right size an operation [416]. No matter the means, good forecasts are vital for the efficient low-cost design of supply networks.

Organisations make use of a variety of forecasting techniques, but largely time-series forecasts, based on historical data [417] as exemplified by the Box-Jenkins methods [418]. Forecasting models range in complexity from naïve methods [419], to more advanced moving average and weighted average techniques [420] and into complex signal processing approaches such as Fourier analysis [421] or more lately, so-called Artificial Intelligence methods [422].

We revisit a model that we have previously proposed [370] that shows the complex construction for demand. Total demand that arrives, is constituted by the real, necessary, and correct first time demand which Seddon refers to as value demand and its counterpart demand modality which is due to prior system failure, namely failure demand.

6 - 4.4. Failure demand

Failure demand originates from a prior interaction in which the service was not delivered to resolution, or as Seddon says, demand that is caused by "not doing something or doing something wrong for the customer" [44 Pg. 76]. Our model from previous work shows this behaviour in Figure 6-28 The critical element for failure demand, is the point of system failure – the point at which the system fails to do something, or the right thing [44]. This is the seed for a variety of possible next steps. Once system failure has occurred, the demand exits the system and may either not return at all as shown by ① or return at a future date to attempt the same service ② or moving to a different hierarchical level (such as moving to a different facility or care level) ③, a mode that we have introduced as failure demand II.

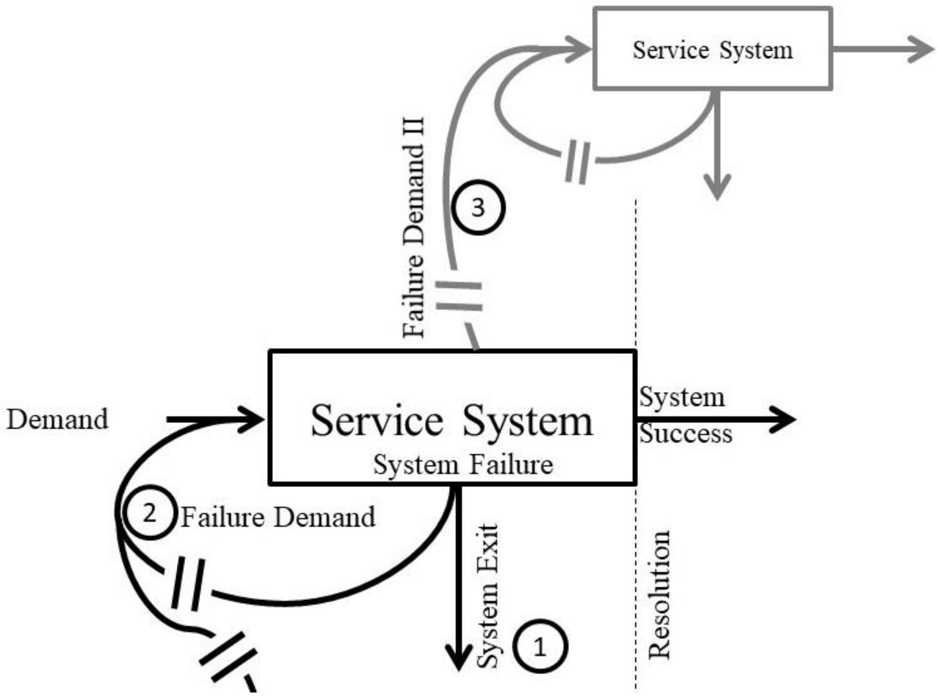


Figure 6-28: Failure demand in hierarchical systems

Failure demand in a system exacerbates the demand that is seen. Because there are repeat interactions for the same service, all the work done on the processing other than the final, successful interaction is waste [63]. Therefore, reducing failure demand directly reduces total demand and improves system efficiency.

The impact of failure demand on service systems has been studied before. In planning departments, between 20 % and 30 % of applications return as failure demand [235], Whilst in banks this was found to range between 40 % and 60 % [110] and in local government call centres, the value could be over 80 % [185].

6 - 4.5. Summary

Measuring supply network performance is important for efficient running and improvement and this measurement can be achieved by a variety of quantitative and qualitative frameworks. An efficient supply-chain network is built to be low cost and as efficient as possible. To achieve this, a good understanding of demand is essential, which can be calculated using a range of forecasting techniques of varying complexity. The system-design should be strongly influenced by these forecasts. Variability is first partitioned to reduce the underlying input signal to a pure one, however forecasts typically do not granularize demand, but merely assess the input signal, which is previous movement data.

We propose that failure demand is a major underlying constituent of total demand and as such, places a considerable burden on the operations of a supply network. Therefore, understanding the impact of failure demand on total demand is of considerable value and should be investigated.

6 - 5. Method

This exploratory study consisted of two case studies [16] both conducted in the same national network, at different levels of the supply-chain. We made use of Visual Basic for Applications to data-mine ordering and shipping records to extract important supply-chain metrics. These are represented making use of descriptive statistics.

Determining failure demand is not simple and unlike most quantitative KPIs (such as On-Time-Delivery, Stock-Turns etc [403]) failure demand is derivative and cannot be measured on the transaction data of a particular interaction. For that reason, a computational approach is required. An occurrence of failure demand only takes place when a previous system failure has led to a repeat interaction (recall Figure 6-28). We had to develop custom code, that would step through the historical data, discover instances of system failure, and then search through the data again, to find an instance of repeat interaction. If in turn this instance was also system failure, the horizon was stretched further.

6 - 5.1. Software development

Visual Basic for Applications (VBA) [423] is a powerful addition to Microsoft products such as MS – Excel. It is a programming language that allows for automation of processing tasks and meta-analysis which cannot be done using traditional means. For this study, approximately 1 200 lines of code were developed to automate the cleaning, reformatting, processing, and analysis of the datasets.

6 - 5.2. Data sources

Several reports were extracted from E-Med which contained logistic information about product movement. We created two datasets from these reports, namely an order database and a dispatch database. The specific data-classes in each are summarised in Table 6-4: :

Table 6-4: Database Classes

Order Database	Dispatch Database
Item Description, Number Required, Health Facility Requesting, Order Date	Item, Unit, Quantity, Facility, Facility Type, Region, Zone, District, Unit Cost, Amount, Sales Record Number, Printed Date, Payment Type

Additional databases were extracted from E-Med for cross referencing and processing, which included:

- Pharmaceutical reference price list – this list relates products to their reference prices
- Drug equivalency lists – this list relates drugs that can be used as pharmacologically equivalent, either on formulation, dosage or due to inconsistent naming or differing manufacturers.
- Inventory snapshots - three snapshots were received per network over the course of the study and were used to determine system-wide inventory levels.
- Drug classification lists – identifying medication as critical, essential or non-essential drugs.

A full dataset as described here was compiled for each network, namely the wholesale-and the distribution -networks.

6 - 5.3. Data validation

The first step in processing the data, was to validate the accuracy of the dataset, and to do so, several sub routines were written. Common errors that had to be checked are shown in Table 6-5.

Table 6-5: Data Validation

Data type and format checks	Data were checked against their expected types and formats. For example, all dates were assessed whether they were stored as dates and whether their format was consistent. In this case, all dates were converted to a consistent yyyy/mm/dd format.
Presence checks	All data rows were checked for missing data points. All such rows were deleted.
Spelling and grammar checks	This assessed the consistent use of descriptors and assessed the extent to which the same item or location was consistently described. This required the creation of equivalency tables and considerable corrections were made
Uniqueness checks	Data points were assessed for uniqueness, and all full duplicates were assessed. In some instances, sites would submit two duplicate orders, as a strategy to improve delivery chances, however in other instances, duplicates were errors. Discretion was used to differentiate between these.
Cardinality check	Cardinality assesses the level of variation within a data column [424] this allows a rational check to assess whether the data spread is realistic. We checked the cardinality for all columns and checked whether they conformed to expectations. For example, high cardinality was expected for product name, and date where major variation existed, whilst for sites the cardinality had to be lower.

6 - 5.4. Data processing

6 - 5.4.1. **Creating a master document.**

The raw order databases are structured per line item, and in the form that the original orders arrived. The first subroutine after cleaning the data, was to recreate the original order (as it arrived) by clustering all orders that arrive from the same customer on the same day. Once the order master had been rebuilt, deliveries were matched to orders to infer order-performance.

In the absence of differentiated order numbers and these reflecting at dispatch, the code stepped through the master document, seeking a matching delivery for an order. The routine matched site, product and date (within ten days tolerance). If a match was found, the delivery was recorded in the master document. Order volumes and delivery volumes were recorded. The products equivalency database was used, as often products were ordered using one descriptor and shipped based on another[28,29], similarly, we also applied a site equivalency table [30]. The master document served as the source for all further analysis.

6 - 5.4.2. **Preliminary metrics**

A selection of basic metrics, as explored in section 6 - 4.2 was assessed to determine general system performance. These metrics accompanied by calculation notes are shown in Table 6-6.

---

[28] "Examination Glove - Powder free-Nitrile, Non Sterile, Large" is the same as "Glove Examination Powder Free Nitrile Non-Sterile, Large".

[29] Inconsistent product codes exacerbated this problem as they could not be relied on to discriminate between product.

[30] For examples of site equivalencies, see Section 6 - 6.1

Table 6-6: Key Performance Metrics and Calculation Notes

Metric	Notes
Delivery-in-full	Percentage of all orders placed shipped in full (short deliveries[31] were counted as conforming)
Order-processing-time	Average time taken to process an order - time between arrival of order and dispatch.
Short-dated-inventory	Percentage of all inventory that has less than sixty days of shelf life remaining
Expired product	Percentage of all inventory that has reached its expiration date
Inventory-turns	The hypothetical number of times the inventory could be entirely replaced in a year. Value of goods moved / Average Inventory Value.
On-the-shelf-availability	Percentage of the master inventory in stock at a snapshot in time.

6 - 5.4.3. **Failure demand related metrics**

The first step to assessing failure demand, is to create a database of system failures (recall Figure 6-28). Stepping through the master document, all instances where the delivery volume was below the volume that was ordered by the customer was transferred into a new system failure database. This includes instances where partial, or no, deliveries are made. This database represents the base case of all system failures.

[31] In this instance a "short delivery" was a delivery that was made however fewer items than ordered were delivered.

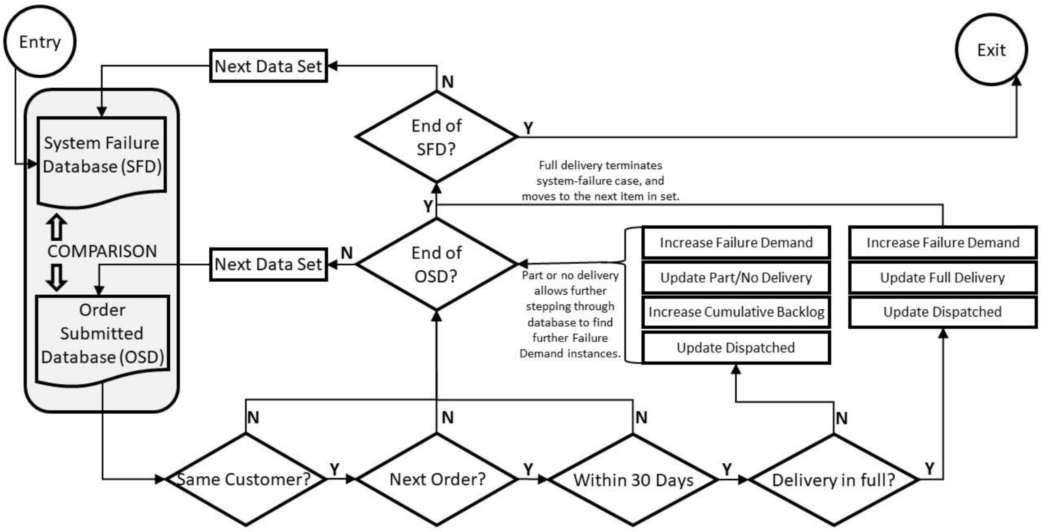


Figure 6-29: Algorithm for determining failure demand developed for this study

The code (algorithm shown in Figure 6-29, enlarged in Appendix 5, Page 339) steps through the system failure database, and if, by the next order, or thirty days (whichever is sooner) the same order repeats from the customer, then this instance of demand, is interpreted as failure demand. If this instance again does not have a satisfactory resolution, then failure demand is assessed for the next event. This loops until the entire database has been compared to the system failure database line – the next line in the system failure database is now assessed.

Data that were assessed for this case included: Cumulative-order-quantity, Cumulative-dispatch-quantity, No-dispatch-count, Partial-dispatch-count, Full-dispatch-count and critically for this study, failure demand-count.

Equation 17 shows a relationship on which basis the prevalence of failure demand can be calculated.

$$\mathrm{FD}_{\%} = \frac{n_{FD}}{n_O} \times 100 \qquad (17)$$

With $FD_{\%}$ as the percentage of failure demand, n_{FD} the count of failure demand instances and n_O as the number of original orders.

Note that the denominator is not only the data in the system failure database, but all orders, regardless of system failure, as failure demand expresses a proportion of total demand.

6 - 6. Results

6 - 6.1. Data validation

E-Med successfully managed the data entered so no duplication errors were found. Data type and format checks showed that the data had been well managed and as a result, no corrections were required. Presence checks showed no evident gaps, again a result of E-Med's gate-keeping role.

The major errors that were found were consistency errors in terminology for site names and product descriptions. To avoid case errors, all labels were converted to upper case and all punctuation was removed. After this step, numerous non-keystroke errors remained.

For example, "Denver[32] Health Centre", "Denver H/C" and "Denver Health Center" were all mapped to the same health-facility as was "New York[32]

[32] This study did not take place in the USA. The locations are used as an example mimicking the actual errors.

Hospital" and "Newyork Hospital". The order of descriptors varied, sometimes being adjective-noun "Nitrile Gloves" or at other times noun-adjective "Gloves, Nitrile". Other common errors were inconsistent applications of US and UK spelling, for example "Disk" versus "Disc". The depth of detail was also inconsistent, with "Heater", "Electric Heater", and "Electric Heater 2000W" all appeared in the datasets. Inconsistent formatting also presented difficulty; for example, the equivalent "Abacavir (ABC) + Lamivudine (3TC) - (600mg + 300mg) - Tablet" and "Abacavir + Lamivudine - (600mg + 300mg) - Tablet", had separate database entries, which was particularly problematic as E-Med did not make use of unique product identifying codes. We created a site mapping tool that had 200[33] corrections and a product mapping database that contained 1 917 equivalencies.

After the data had been cleaned satisfactorily, we were able to process the data to recreate the original orders.

6 - 6.2. Summary of data processing

From the raw data, the master ordering- and dispatch-databases were created, and the information is shown in Table 6-7. As can be seen, the data sets reflect approximately a year and a half of operations at both sites. The wholesale network may have fewer order lines; however, this is compensated for by substantially greater volumes. The disparity between Order-lines-delivered and Product-lines-delivered[34], is because the hubs often push material [15] to their customers despite an absence of a request.

[33] This number is accurate and purely coincidentally round.
[34] An order line is a line item dispatched on request of the customer, a product line delivered does not correspond to a customer order yet is shipped anyway.

Such practice is not Lean, and inefficient, leading to higher cost, redundancy and generally poor performance against KPIs.

Table 6-7: Database Information

Database Characteristics	**Wholesale Network**		**Distribution Network**	
	n	**%**	***n***	**%**
Days in period	513	Ref	512	Ref
Days on which orders were received	455	88.7%	354	69.1%
Days on which orders were dispatched	348	67.8%	339	66.2%
Order Database Lines	92 745	Ref	202 003	Ref
Dispatch Database Lines	10 034	10.8%	106 419	52.7%
Orders Delivered	11 500	12.4%	75 997	37.6%

A selection of traditional KPIs is shown in Table 6-8. It is clear from this table that the system is underperforming, with Deliveries-in-full under 40 % and more worryingly, in the case of the wholesale network, under 10 %. The remaining KPIs have considerable room for improvement, though some positives include the surprisingly low short-dated product and the unexpectedly high number of stock turns per annum. Many of the problems in the supply-chain can be traced back to very low on-the-shelf availability of stock, which is an area for policy intervention.

Table 6-8: Performance as measured against KPIs [403, 409]

	Wholesale Network		**Distribution Network**	
Traditional KPIs [403]	***n***	**%**	***n***	**%**
Delivery in Full (%)	7 132.0	7.7%	74 337.0	36.8%
Order Processing Time (days)	1.9		2.9	
Short Dated Inventory (%)	0.6		0.1	
Inventory Turns (n / year)	5.2		5.8	
On the shelf availability (%)	9.3		17.5	

The properties of the system failure database are shown in Table 6-9. Note that for both cases, the prevalence of system failure in other words, the system not satisfying the needs of the customer, is above 70 %. Note that undelivered material is over 60 % in both cases, and nearly 90 % for the case of the wholesale network. The wholesale network was requested to ship roughly nine billion items, yet only managed to ship 43 Million, an efficiency of virtually zero. The distribution hub fared slightly better, with an ultimate efficiency of about 19 %.

Table 6-9: System failure Database Properties

	Wholesale Network		**Distribution Network**	
System failure Performance	***n***	**%**	***n***	**%**
System failure - Lines	87 786	94.7%	149 632	74.1%
Delivered in full	4 959	5.3%	52 371	25.9%
Delivered in part	6 541	7.1%	23 626	11.7%
Not delivered	81 245	87.6%	126 006	62.4%
Total Items Ordered ('000)	9 566	Ref	34	Ref
Total Items Shipped ('000)	43	0.5%	6	18.8%

Based on the data presented in Table 6-9 and using the logic derived earlier, this means that all instances where the orders were not satisfied in full amount to system failure.

From the system failure database, return interactions corresponded to a prior system failure were recorded as failure demand. This is shown in Table 6-10. Note that the full deliveries remain low – however a full delivery terminates a previous system failure. Therefore, in the wholesale network, only 2 846 instances of system failure were terminated by the supplier, which amounts to 3 % of system failure. The same number for the distribution network was 7.2 %.

Table 6-10: failure demand Results

Failure demand Measures	Wholesale Network *n*	%	Distribution Network *n*	%
Orders (O)	92 745	Ref	202 003	Ref
System failure (S)	87 786	Ref	149 632	Ref
System failures Resolved (/S)	2 846	3.2%	10 773	7.2%
Failure demand (/O)	51 623	55.7%	58 615	29.0%

Importantly for this study, using Equation 4, we discovered a failure demand value in a health care setting, which is 56 % in the wholesale network and 29 % in the distribution network.

6 - 7. Analysis and Discussion

Shortages and non-delivery (which we think of as system failure) are common issues that are monitored in supply-chains. And although the impact of these is generally appreciated, the derivative measure of failure demand is not broadly described nor is its impact on system performance appreciated, neither in terms of capacity nor in terms of demand.

6 - 7.1. Failure demand

The failure demand experienced in the distribution network (29.02 %) is considerably lower than the failure demand experienced in the wholesale network (55.66 %). There are several reasons for this difference based on operational and other reasons. Recall that failure demand is the phenomenon that happens *in response* to an initial system failure – in effect – customers who return for the same service. There are good reasons why customers are more likely to return to the main hub compared to the secondary hubs, these include proximity to the need, access to the free market and order volumes.

The customer to the main hub (wholesale) is the secondary hub whilst the customer to the secondary hub (distribution), is the health facility. This is an important distinction, because the health facilities directly engage the patients and have the most critical need for pharmaceuticals. By contrast, the hubs are warehouses, and as such the sense of urgency, when faced with shortages, is absent. This may suggest that failure demand at the secondary hub should be higher than at the main hub, however, the health facilities, when encountering system failure are more likely to go to the open market to procure product, driving down apparent failure demand in the distribution network. By contrast, the secondary hub may be able to absorb shortages, and thus avoid going to the open market – thus resulting in higher apparent failure demand here.

Accessing the open market is also easier in the case of health facilities, as they tend to place significantly smaller orders, meaning that they can absorb higher prices but also do not need a vendor capable of supplying significant volumes. The secondary hub however generally places very large orders for which the price premium on the open market may be prohibitive whilst the availability of product in the required volumes is rarely available in-country, and certainly not at short notice.

Health facilities have more direct and more practical access to the open market for procuring shortfalls. This means that even though the shortage results in system failure, they do not return to the same location but procure product outside of the formal supply-chain. Consequently, they do not reflect as failure demand, however, it is our claim that they still constitute failure demand if we treat the entirety of the supply-chain as the system. Unfortunately, we had no visibility of such transactions and can merely deduce their presence and magnitude. We therefore conclude that the failure demand rates that we report are conservative for both systems, but particularly for the distribution network.

Processing shortage letters places a heavy administrative burden on the system. A shortage letter is drafted for every shortage that is recorded, and the letter is issued to the customer. The time taken to compile such a letter includes the (non-automated) preparation of the document, printing, and having it authorised by the signature of the branch manager. The process on average takes 2.1 minutes to complete per order line. This process added over 5 000 hours of labour over the course of this project, amounting to over fifteen work-hours per day, effectively requiring two staff members dedicated to the task of dealing with the paperwork for failure demand. This number of staff was verified in actual observations.

6 - 7.2. Travel time and distance

An additional burden on the system is on the physical infrastructure. Due to failure demand, health facilities need to make repeated unnecessary trips to the hubs. This results in time loss for drivers and pharmacists and additional wear and tear on vehicles, raised fuel and other costs and an associated risk cost due to travelling under frequently hazardous conditions.

6 - 7.3. Reservation of pharmacist time

Orders will not be dispatched from the secondary hub unless a competent representative from the site – generally a pharmacist – is physically present to sign off the order. The average time taken from order submission to order sign off is 2.4 days. This means that for smaller facilities, during this time, the pharmacy may be inoperative, because the (often only) pharmacist is away at the hub, or not competently managed. If failure demand requires the repeated presence of a pharmacist, this number is compounded, resulting in compromised downstream service delivery and care at health facilities.

6 - 7.4. People and orders in queue

A practical way to interpret the failure demand values, is to interpret them as opportunities for reduced demand. In effect, a 56% failure demand means that 56 % of the administrative work done at the main hub is done more than once, and that initial work efforts were not value adding. This is over and above the resources dedicated to managing the administration when system failures occur. Given that adding small amounts of load to a system with high utilisation can make the queue length go catastrophically out of control [39, 43]. This means that small amounts of unnecessary demand can have a non-linear effect on the system's capacity [194]. The inverse is also likely true, i.e., that when failure demand is removed, the net gains can be even greater than the actual magnitude of the reduced load.

6 - 7.5. Achieving gains

Seddon emphasises that failure demand should be measured not for the sake of target setting [44] but rather as a means for identifying poor processes and systems that can be improved. To make any gains in this domain requires system wide scrutiny [137, 320]. The Vanguard technique says that Failure demand must not be measured on an ongoing basis, but rather as a diagnostic part of improvement interventions [32]. Therefore, we take our measurements as a snapshot of the system which can be used to identify opportunities for improvement. Using the principles of Lean to manage operational issues, but using the philosophically powerful concepts of Lean thinking, as advocated by Hines et al. [174] to design the future shape of the system. We further emphasise that simply because gains are possible, this does not mean that achieving them is easy or even achievable [210].

To achieve any gains, some interventions are required from policy through training to practice. We realise that measuring failure demand is easier than managing it, however, we do believe that awareness of failure demand levers is important for streamlining this pharmaceutical supply-chain; reducing overall operating cost, increasing performance and reducing shortages and improving patient wellbeing [305].

One of the major drivers of failure demand, was the high level of stock-outs. We attribute this to inefficient procurement policy, which focuses on infrequent orders of very large batches, with associated losses due to pilferage and expiration. This also means that if there is a system wide stock out, replenishment can take an extended time. Lean ordering principles and moving towards a pull system, rather than the current push [63] would meaningfully improve the efficiency of the supply-chain. This should enable lower stock-outs and generally improve system behaviour, leading to improved customer service.

The idea of "learning to see" suggests that an important means of understanding and solving problems is in developing an ability to see them. Many problems are invisible until people know what to look for [425]. This notion of learning to see, is central to Lean, where understanding which wastes to look out for in a system allows practitioners to identify problematic areas and opportunities for improvement [65]. Richardson speaks of the philosophy to "go see" at Toyota, where a well calibrated eye, trained to see waste, variation and problems discovers opportunities for improvement [426] Similarly, we believe that failure demand is an opportunity for "learning to see". Whilst many contextual issues do need resolution, creating a cohort of people who understand the effect of failure demand and who are able to recognise both system failure as well as failure demand, will innovate strategies to limit its impact.

6 - 7.6. Limitations of this study

This study was conducted in a single developing country pharmaceutical supply-chain. The systemic issues related to stock shortages certainly exacerbate the failure demand in the system.

6 - 8. Conclusions and Recommendations

Our framework can be used as the underlying guideline to identify system failure and failure demand in health settings.

We have developed an analytical tool that uses order- and dispatch-data to extract a system wide failure demand performance indicator. The tool works and has been tested on two sites within the same health system. The tool that was developed shows that simple, commonly available software can be used to develop relatively sophisticated data mining tools from first principles, without the need for costly data mining software that may be out of reach for researchers and practitioners in resource constrained environments.

We recommend that this same tool is tested in other developing pharmaceutical markets and extended into developed markets, to test our belief that failure demand is not a problem confined to developing economies.

Failure demand was calculated for two relationships – for wholesale network (55.66 % of all demand) and the distribution network (29.02 %). Our failure demand values are range validated, being of an expected magnitude, based on existing work [110, 185, 235]

The failure demand numbers can be interpreted to be the extent to which the load is greater than it should be. The additional load the systems likely has a non-linear effect on total load, meaning that demand can be reduced by more than the amount of failure demand identified.

Additionally, the resources dedicated to managing failure demand (through the creation of shortage letters, other administrative processes and the re-formulation of orders) at the hubs and sites, significant labour reductions are possible with no productivity losses.

We believe that both values that we present are conservative, because customers facing system failure also have remedies on the open market to procure product that was unavailable in the formal supply-chain. We believe that this effect is even more significant in the distribution network due to the health facilities' proximity to the patient.

Failure demand introduces a diversity of performance issues to the system. Policy requires a pharmacist from site to sign off and accept the order at the hub. The average order processing time of roughly two days, means that pharmacists are off duty during this time, and in the case of failure demand, the negative impact on the site is compounded. Other issues include that the open market may be more costly, placing a more complex sourcing burden on pharmacists at sites with complex lead times, and may require far greater distances to be travelled.

Vehicles that make journeys entirely to service failure demand results in further costs due to wear and tear and fuel and driver time. In this case, the country is large and road infrastructure poor, so as a result the burden placed on the system by this is significant.

This study contributes towards creating an awareness for the impact of failure demand in a basic hierarchical pharmaceutical supply-chain. The magnitude of failure demand in the system is of concern and improving matters will require, on the one hand, awareness – so that the impact on the system is recognised, and on the other hand, procurement policy change, as infrequent batch purchasing leads to frequent stock outs and a low level of stock availability.

We believe that failure demand is a major contributor to low levels of service delivery in the developing country that was investigated. Managing failure demand can lead to improved service delivery, improvement in the plight of vulnerable people, reduction of cost and can lead to higher efficiency and quality.

6 - 9. References

15. Liker, J.K., *The Toyota Way*. 2004, McGraw Hill, New York
16. Yin, R.K., *Case Study Research: Design and Methods*. 1987, Sage Publications, Inc., Beverley Hills
32. Seddon, J., *The Vanguard Method.* Introduction to Vanguard - eLearning, 2020(1).
39. Bicheno, J., *Queues!*, D. Hartmann, Editor. 2019, Buckingham, United Kingdom.
43. Roser, C., *A Small Dice Game for the Kingman Formula*, in *All About Lean*, C. Roser, Editor. 2017.
44. Seddon, J., *Freedom from Command and Control: Rethinking Management for Lean Service*. 2005, CRC press, Boca Raton, Florida
50. Goldratt, E.M., *The Goal: A Process of Ongoing Improvement,* . 2004, Routledge, New York City
63. Bicheno, J. and Holweg, M., *The Lean Toolbox: A Handbook for Lean Transformations*. 5th Edition ed. 2016, PICSIE Books, Buckingham, UK
65. Womack, J.P., Jones, D.T., and Roos, D., *The Machine That Changed the World: How Lean Production Revolutionized the Global Car Wars*. 1991, Simon & Schuster, London
100. Ackoff, R.L., *Management Misinformation Systems.* Management Science, 1967. 14(4), p. B-147-B-156.
110. Piercy, N. and Rich, N., *Lean Transformation in the Pure Service Environment: The Case of the Call Service Centre.* International Journal of Operations & Production Management, 2009. 29(1), p. 54-76.
137. Seddon, J., O'donovan, B., and Zokaei, K., *Rethinking Lean Service*, in *Service Design and Delivery*. 2011, Springer. p. 41-60.
174. Hines, P., Holweg, M., and Rich, N., *Learning to Evolve: A Review of Contemporary Lean Thinking.* International Journal of Operations & Production Management, 2004. 24(10), p. 994-1011.
185. Marr, B. and Neely, A., *Managing and Measuring for Value: The Case of Call Centre Performance*. 2004, Cranfield School of Management and Fujitsu, Cranfield, UK.
194. Walley, P., Found, P., and Williams, S., *Failure Demand: A Concept Evaluation in Uk Primary Care.* International Journal of Health Care Quality Assurance, 2019. 32(1), p. 21-33.

210. Jack, E.P. and Powers, T.L., *A Review and Synthesis of Demand Management, Capacity Management and Performance in Health-Care Services.* International Journal of Management Reviews, 2009. 11(2), p. 149-174.
235. Seddon, J. and Brand, C., *Debate: Systems Thinking and Public Sector Performance.* 2008.
294. Shanmuganathan, T., *Ethics and the Observer's Paradox.* relc Journal, 2005. 36(1), p. 73-84.
305. World Health Organization, *Constitution*. 1989, World Health Organization, Geneva, Switzerland
320. Seddon, J. and Caulkin, S., *Systems Thinking, Lean Production and Action Learning.* Action Learning: Research and Practice, 2007. 4(1), p. 9-24.
370. Author Withheld, *Reference Withheld for Reviewing Purposes.* 2020.
398. Drucker, P., *The Effective Executive: The Definitive Guide to Getting the Right Things Done*. 2006, Collins, New York
399. Bauer, K., *Kpis: Avoiding the Threshold Mcguffins.* Information Management, 2005. 15(4), p. 47.
400. Muller, J.Z., *The Tyranny of Metrics*. 2018, Princeton University Press
401. Goodhart, C.a.E., *Monetary Theory and Practice: The Uk-Experience*. 1983, Macmillan International Higher Education
402. Strathern, M., *'Improving Ratings': Audit in the British University System.* European Review, 1997. 5(3), p. 305-321.
403. Supply Chain Council, *Supply Chain Operations Reference Model.* Overview of SCOR version, 2008. 5(0).
404. Krauth, E., Moonen, H., Popova, V., and Schut, M. *Performance Indicators in Logistics Service Provision and Warehouse Management–a Literature Review and Framework*. in *Euroma international conference*. 2005.
405. Keebler, J.S., Manradt, K.B., Durtsche, D.A., and Ledyard, D.M., *Keeping Score: Measuring the Business Value of Logistics in the Supply Chain*. 1999, Council for Logistics Management, Boston, Massachusetts
406. Slack, N., *The Manufacturing Advantage*. 1991, Mercury, London
407. Da Silva Moitinho, M., Das Chagas, H.X., Matos, M.C.P., De Oliveira, R.M.N., De Souza, J.F., and De Moura, V.A. *The Use of Indicators to Measure the Performance of Logistics Services in a Steel Industry*. in *PONSMeeting*. Brasilia, Brazil.

408. Fawcett, S.E. and Cooper, M.B., *Logistics Performance Measurement and Customer Success.* Industrial Marketing Management, 1998. 27(4), p. 341-357.
409. Coyle, J.J., Langley, C.J., Novack, R.A., and Gibson, B., *Supply Chain Management: A Logistics Perspective*. 2016, Nelson Education, Toronto, Canada
410. Pycraft, M., Singh, H., Phihlela, K., Slack, N., Chambers, S., and Johnston, R., *Operations Management: Global and Southern African Perspectives.* Cape Town: Pearson Education, 2010.
411. Heizer, J. and Render, B., *Principles of Operations Management*. 2011, Pearson, London
412. Huan, S.H., Sheoran, S.K., and Wang, G., *A Review and Analysis of Supply Chain Operations Reference (Scor) Model.* Supply Chain Management: An International Journal, 2004. 9(1), p. 23-29.
413. World Health Organization, *Annex 5 Who Good Distribution Practices for Pharmaceutical Products.* 2010.
414. Monden, Y. and Lee, J., *How a Japanese Auto Maker Reduces Costs.* Management Accounting (USA), 1993. 75(2), p. 22-27.
415. Rushton, A., Croucher, P., and Baker, P., *The Handbook of Logistics and Distribution Management*. 5th Ed ed. 2019, KoganPage, London
416. Aghezzaf, E., *Capacity Planning and Warehouse Location in Supply Chains with Uncertain Demands.* Journal of the Operational Research Society, 2005. 56(4), p. 453-462.
417. Jain, C.L., *Benchmarking Forecasting Models.* The Journal of Business Forecasting, 2006. 25(4), p. 14.
418. Box, G.E., Jenkins, G.M., Reinsel, G.C., and Ljung, G.M., *Time Series Analysis: Forecasting and Control*. 2015, John Wiley & Sons, Hoboken, New Jersey
419. Hyndman, R.J. and Koehler, A.B., *Another Look at Measures of Forecast Accuracy.* International Journal of Forecasting, 2006. 22(4), p. 679-688.
420. Wheelwright, S. and Makridakis, S., *Forecasting Methods for Managers.* 1974.
421. Fumi, A., Pepe, A., Scarabotti, L., and Schiraldi, M.M., *Fourier Analysis for Demand Forecasting in a Fashion Company.* International Journal of Engineering Business Management, 2013. 5(Godište 2013), p. 5-30.

422. Khalil Zadeh, N., Sepehri, M.M., and Farvaresh, H., *Intelligent Sales Prediction for Pharmaceutical Distribution Companies: A Data Mining Based Approach.* Mathematical Problems in Engineering, 2014. 2014.
423. Albright, S.C., *Vba for Modelers: Developing Decision Support Systems with Microsoft Excel*. 4th ed. 2011, Cengage Learning, Inc;
424. Chassaing, P. and Gerin, L., *Efficient Estimation of the Cardinality of Large Data Sets.* arXiv Preprint Math/0701347, 2007.
425. Rother, M. and Shook, J., *Learning to See: Value Stream Mapping to Add Value and Eliminate Muda*. 2003, Lean Enterprise Institute
426. Richardson, T. and Richardson, E., *The Toyota Engagement Equation: How to Understand and Implement Continuous Improvement Thinking in Any Organization*. 1st edition ed. 2017, McGraw-Hill Education, New York City

Chapter Seven: **Unifying Narrative.**

7 - 1. Introduction

When this study started the aim was to understand how failure demand shapes demand in health systems. To address this, a few fundamental questions had to be answered. What is demand in health care in the first place, where does it come from, what makes it occur, and how bad is it?

7 - 1.1. Positioning of the study

It is difficult to place this study within a dominant philosophy; however, Lean is a logical option for it to reside. Ultimately, the goal is to satisfy the customers' needs, about value, about minimizing waste, about levelling unevenness and improving overburden. Yet, John Seddon, the father of failure demand is highly critical of Lean, seeing it as a codification of method [69], proposing that Lean is a "waning fad" [138] and is very disparaging of people whom he thinks of as "tool heads" [119]. Tool heads are people who discard the philosophical underpinnings of Lean in preference for structured tools that emerged when American business leaders visited Japan after Deming's time there [44]. Similarly, the study also fits within systems thinking, which satisfies the idea of "Joined-up-ness" [101, 235] and Seddon also proposes that systems thinking forms the underpinning for Lean [320].

Our lens is strongly shaped by the Lean philosophy (and not as tool heads), however we prefer in this case to frame our contribution in terms of an aim, rather than in terms of a process. Therefore, rather than polarising the outcomes of this study by framing it in a single philosophy, we prefer to position this study in the broad domain of *Health Systems Strengthening*, a priority of the World Health Organization [427], and the means by which we achieve this, is governed by our broad view of what Lean is, and what it does.

Thinking of the health provision in a systems-way is key to understanding the demand burden. The health system is complex [428], and includes a multitude of actors [96]; hospitals, clinics, pharmacies supply-chains, nurses doctors, porters, patients, taxis, ambulances, scalpels, beds, computers and their systems, money, records, patient histories and a lot more. If all of these do not work together, creating an emergent entity [97], with the aim of delivering health to those in need of it, then system failure of some sort occurs. And the patient is not exonerated from a role in system failure – this is important. It is not a vector pointing from the service provider to the service recipient; that the system achieves its outcomes is a partnership.

7 - 1.2. The ping pong game of blame

In Seddon's words, failure demand is the demand caused by "the failure to do something or to do something right" [44, p.76]. As was discussed in Section 2 - 9.2.3, it is emphasised that the burden of doing things right lies equally on the doctors and the hospital and the supply-chain as it does on those who receive the service. For system failure, it is important to be cautious to not think in terms of "fault" but rather system design [15], however for this example, "fault" will caricaturise a familiar exchange; consider the following:

A patient is denied care at a diabetes clinic, this is the fault of the clinic- however, the patient arrived on the wrong day, the fault is now that of the patient. The clinic however did not make sure to communicate the correct date to the patient – clinic at fault again. But the patient has agency, the patient should have made sure to find out – patient at fault. But the clinic did not have a system in place to make finding out easy; phones just ring, no website, no text messages – clinic's fault. But the patient knows these frailties of the clinic, and should have doubly checked, in person at the last visit – patient fault…

This ping pong game can continue – and does. Service providers generally stop the argumentation when there is suitable evidence that the patient is at fault and similarly patients generally stop when the blame can rest with the service provider. In a Lean organisation however, failure is a great opportunity to improve system design [63]. Therefore, Seddon's formulation should be carefully considered before automatically assigning "blame" to the service provider. Indeed, the purpose of such an investigation should always be to identify systemic design issues and improve them [184]. For example, in this ping pong game, simple interventions such as text messages, a phone app or even an updated website could reduce system failure considerably. That however will only happen, if the pursuit is for solutions, rather than for finding the party at fault.

7 - 2. Demand

7 - 2.1. Population model

The first contribution made in this thesis was to define the demand population for a health system (see Figure 4-20 on Page 171). This consisted of defining the idea of what the unhealthy population is, classifying that as the natural demand population for the health system, but recognizing that not all the unhealthy population becomes load in the system. And furthermore, realising that some of the healthy population too, forms a burden on the system.

7 - 2.2. Demand modalities

A model is proposed that modularises demand in health more granularly, and breaks it up into value demand and failure demand (Seddon's categories [44]) but goes on to add the important classifications of escalation demand and false demand. These new modalities complete the range of load that a system experiences.

This modularisation is structured by systemic mechanisms compared to Vickery & Lynch' structure that views demand through a lens shaped by patient-motive [19]. Figure 7-30 shows that there is not a one-to-one correspondence of Vickery & Lynch' classification system [19] to the one presented in this thesis. For example, morbidity demand can mechanistically be value demand, failure demand as well as escalation demand. For that reason, Vickery & Lynch' model makes it difficult to manage demand as the root causes for the different modes in their model would be different systemic drivers. Similarly, apart from morbidity demand, all other modes can be interpreted as false demand, in other words, demand originating in the healthy population.

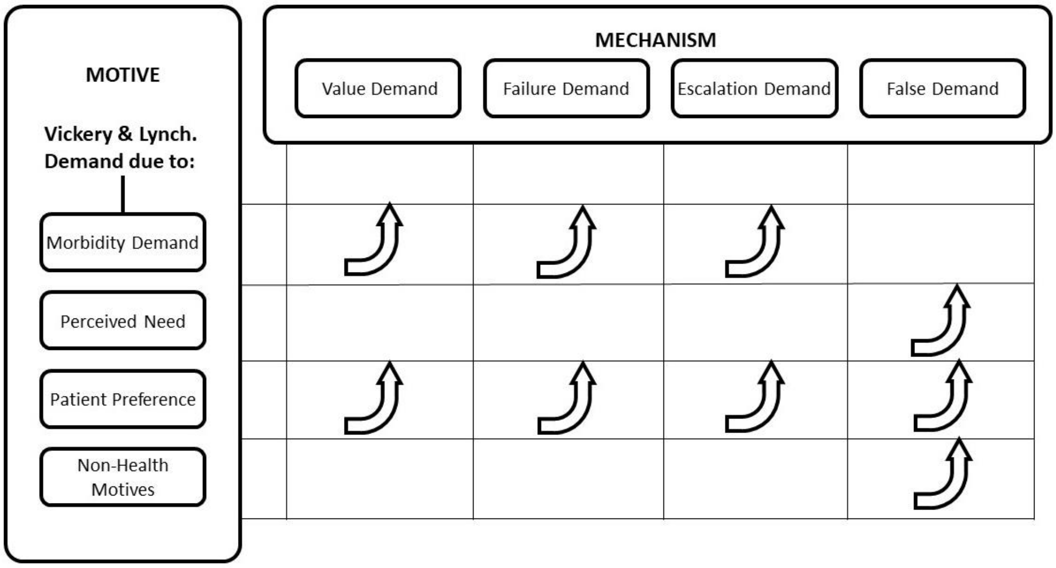


Figure 7-30: Comparison of Motives and Mechanisms of Demand

The model of this thesis shows the mechanisms that shape demand. It is therefore proposed that it presents a rational means of assessing demand constitution and allows for good root cause analyses in system redesign. As such it is a useful granular tool that can strengthen a health system by classifying demand according to the mechanisms that have created it, rather than by the motives underlying care seeking behaviour or the patterns that are stripped of underlying drivers. The key driver in this case is system failure.

7 - 2.3. System failure

Defining failure demand is easy; it is merely system failures that return [44]. Defining system failure however is complex, requiring systemic understanding of the operating environment to judge whether indeed the system has failed, and if so, how.

Several authors have alluded to the idea of system failure, among them Piercy & Rich [110] who speak of service failure, and Masters who identifies events that induce demand [240]. This thesis however emphasises the importance of differentiating language to ensure that failure demand and system failure are not conflated.

To do this, an algorithm was constructed that is able to assess whether a failure can be classed as a system failure (in other words, related to system design) or whether it falls in the 6 % of Deming's 94 – 6 rule [239]. This algorithm was designed to assess system failure in the generic case in a complex hierarchical organisation, such as health care.

The algorithm created a mental model defining the specific case in health systems and arrived at a system failure go- no-go- gate, describing four high level domains in the health system that could seed failure demand. These include hierarchical movements in the system, unsuccessful medicine, patient errors and the operational environment. Each of these contain further sub-categories.

7 - 2.4. Failure demand

One of these sub-categories – "poor supply-chain and inventory management" was chosen to test the impact on a selected domain in the health system, due to failure demand. The choice of the element was due to data availability and other convenience factors. Further studies in the other verticals would be useful.

Because failure demand is more complex to assess than standard KPIs, customised data mining code had to be created. This operated on a Sequential Pattern Mining (SPM) principle [429][35] that steps through the databases and firstly finds system failures, based on the model, and then reinvestigate the database to assess whether the cases returned for the same product within a reasonable time. This led to what Mabroukeh et. al. call "suffix growth" on a database projection [430].

This thesis has delivered useful results and shows that in the pharmaceutical supply-chain of the health system, service delivery is compromised by the low level of stock availability and poor system planning.

The low performance and the high levels of failure demand drive several unwanted behaviours, which may include stock hoarding or hedging strategies at warehouses and at sites, which would further imbalance the system, with over supply at one place competing with stockouts elsewhere.

[35] This also showed that data mining from first principles is possible and is particularly useful for researchers in more resource constrained environments. Access to generally available software opens the possibility for such researchers to do research that would otherwise be constrained by a lack of access to sophisticated data-mining software.

7 - 2.5. Agency

Throughout this thesis, the terms motivation and mechanism were used to describe different points of view for constructing mental models that define the building blocks of demand.

This can also be interpreted from the point of view of the main actor, and with whom the agency for demand construction lies. From the point of view of this thesis – demand is shaped by the actions of the system – which includes patients. Demand is also shaped by patient motivation, but this thesis proposes that managing patient behaviours and their motivation, is also a function of the system. Meaning that in both cases, the motivation model of Vickery & Lynch. is subordinate to the system driven model presented throughout this thesis.

Regardless, this thesis believes that Vickery & Lynch' model is a useful model for describing the agency of the patient, whilst the formulation proposed in this thesis represents a model for the agency of the system.

7 - 2.6. **Cautions about results**

Seddon highlights that failure demand should be used as a snapshot and indeed as an occasional analysis to identify opportunities for improvement [44], a view that is widely supported [110, 185, 194]. This is generally because the alternative is to treat failure demand as an indicator, which leads to the ingenuity of managers being applied to finding ways to report better numbers instead of applying ingenuity to identifying and improving the root causes of failure demand [44].

Failure demand however is frequently used inappropriately as a KPI and it is often forgotten that failure demand is merely the first step of the Vanguard method [32]. Following on from understanding demand it then becomes important to redesign the system to eliminate the root causes of failure demand and that this is no trivial task.

7 - 2.6.1. **Gains**

Fixing the underlying issues is complex. Failure demand cannot simply be deleted by recognising that it exists. It is important to distinguish between theoretical gains, which could be the full amount of failure demand and the gains constrained by practical considerations [44, 110, 185, 194]. This is particularly important as unmoderated expectations are likely to lead to disappointment, thus reducing future support for improvement projects as was shown in the work presented in Chapter 2 [111, 142].

For example, Piercy & Rich [110] show three nine-month case studies in which Lean interventions were made. Tracking important metrics, they reported improvements across the board. The range of failure demand reduction was between 36 % and 68 %. Simultaneously, there was no linear pattern between failure demand and other metrics that were reported. This is characteristic of the non-linearity of systems [92, 105].

The key caution here, is that even though failure demand can be used as a powerful change-lever [44], expectations must be moderated. It must not be expected that the full amount of failure demand can be eliminated.

7 - 2.6.2. **Gains will most likely be lower**

The amount of failure demand in a system can theoretically be the amount of load on the system, that the system should not have to deal with. However, practically, it is an error to believe that all failure demand can simply be deleted, leaving the system unburdened to the full extent of the load placed on the system by failure demand.

Practically, implementing change is often difficult. In the case of call centres, most of the failure demand was a consequence of the "command and control" metrics (and reward programs) commonly in use [110, 185]. These induced call operators to conclude calls quickly (as rewards were based on calls being ended within a target time) and to trick the tracking system into believing that callers had been holding for less time than they had (by answering and immediately returning callers to the queue) [185]. When the focus moved to equipping call centre agents to solve problems and ensure first-call resolution of problems, the improvement in all performance metrics was very impressive [185]. Even so, the gains made in eliminating failure demand were not complete – only about half of failure demand was eliminated [110].

One reason why failure demand cannot always be removed is because it may be woven into the way that systems operate. Masters introduced the term "Potentially Avoidable Failure Induced Demand – (PAFID)" [240] to refer to a phenomenon with major similarity to Seddon's failure demand [44]. Masters shows that a significant portion of PAFID can (currently) not be eliminated. This includes banter and small talk, history taking and other functions that are not truly value demand, but are not avoidable either. This mimics the thinking around necessary-non-value-adding activities [63]. Some activities may be non-value-adding (a good example is taking clinical histories). Just as any other necessary but non-value-adding activities, these cannot simply be eliminated without placing the system at risk.

7 - 2.6.3. **Gains may be higher**

Although this is probably the less common case, theoretically the gains brought on by a focus on failure demand could exceed the amount of failure demand identified.

Learning to see, leads to learning to see more. Green says that to make more evidence-based decisions, evidence needs to be more practice-based [431]. Rother speaks of the idea of learning to see [425]. His idea addresses the fact that people are often unable to see phenomena around them until their attention is drawn to these. Although Rother proposes the tool of the Value Stream Map (VSM) to do so – the underlying principle of teaching people to see is important. Richardson and Richardson's exhortation to "go see" exemplifies the Lean problem-solving mindset. They point to understanding by observing and asking many questions; knowing how to see – how to recognise problems, as the first step of solving problems [426].

The system dynamic archetype of "success to the successful" [99] suggests that in some systems; success begets success. Similarly, in improvement initiatives, the more initiatives take place, the more adepts people become at improvement, see greater opportunities for improvement with greater ease, and execute these more elegantly and swiftly with better outcomes. If failure demand can be the driver for this growth, then gains from finding failure demand may be greater than expected.

Recall the Kingman-Bicheno graph (Figure 2-3 on Page 8). This shows how at higher utilisation, the system load rises exponentially, and small changes in utilisation can cause the system to go out of control rapidly. The curve can be used to show how a small increase in utilisation can have dramatic effects in terms of unmet need, but this also works in the opposite direction. In other words, a small utilisation-shift decrease (for example due to the reduction of failure demand) could lead to non-linear improvements in the overall queue length, which may have system-wide impacts. For this reason, especially near absolute utilisation, the benefits due to removing failure demand from the system may be more meaningful than expected.

7 - 2.6.4. **Caution about hardwiring**

One of the questions asked by the expert panel in Chapter 5 was "Should any distance from the ideal be considered to be system failure / failure demand". This is a good example of where the difference between failure demand as a KPI and failure demand as a means to trigger improvement becomes clear [44].

If system failure can be used to illuminate the distance to the ideal, then work can be done on moving the reality closer to the ideal. If the assessment of system failure is not made, then the gap is not highlighted and the power of failure demand to be a means to system improvement is lost. This aligns with good Lean practice of calling activities "non-value adding" when in doubt [14, 15, 63, 65]. The opposite approach, being hesitant to identify failure demand, obviates the good that can be done by scrutinising the system. In this case, activities will most likely be permanently viewed as value adding [63] (or not failure demand). Therefore, even if no current remedy can be designed, it is important to label all distances to the ideal as failure demand.

It follows that the above will not work if managers are under pressure to report better failure demand numbers. Under such circumstances, it will be unfair to report on numbers over which they have no control, and may lead to resentment and gaming or neglect, leading to the eventual abandonment of the measure (consider the case of NI014 as described in Section 2 - 9.2.4.1).

It is important though to re-evaluate practices. For example, even though chronic care is not failure demand (as was proposed in Chapter 5), this does present opportunities for system improvement and innovation. For example the frequency of some "care" interventions [241] can be reduced, for example through self-testing for stable diabetic patients, thus freeing up system capacity. Therefore, something that is not currently thought of as system failure should not be protected from future scrutiny, and potential reclassification and innovation.

7 - 3. Implications of this Research

The research presented in this thesis suggests that failure demand has a considerable impact on service delivery. It is proposed that the impact of failure demand exists in the practical, operational domain, but that it also has implications in terms of policy, training and society.

7 - 4. Notes on implementation

Lean and failure demand have an uncomfortable history in public services [111, 186, 192, 194, 210]. Initiatives fail for several reasons and these are varied and diverse [30, 111, 142]. Conscious of this, this section attempts to show the areas where caution is important to ensure that the value of the work is not lost in an implementation with unreasonable expectations of possible gains.

In the context of this thesis, Chapter 6 showed that failure demand amounted to as much as 56 % of total workload on the system. This number however must be interpreted with caution. It must firstly not be assumed that the full value of failure demand can be deleted from the system, nor, secondly, that reducing the failure demand by any amount is easy.

7 - 4.1. In this case:

The root causes that emerged from analysing the failure demand in the system included issues associated with the supply-chain strategy. This was largely driven by very long lead times (sometimes more than 14 months) which encouraged the procurement of large batches, which were infrequently delivered. As a result of the periodic shortages, hubs as well as health facilities over-ordered and hoarded for the anticipated gaps in supply. This was exacerbated by inadequate supply-network-, contract-, and donor-management. Which in turn was aggravated by poor forecasts that made use of untrustworthy consumption data, due to high hoarding levels and ironically very high obsolescence of products.

Needless to say, these issues cannot be simply resolved, but require system-wide strategic interventions [174], to drive down lead times and batch sizes, thus stabilising system wide supply-dynamics. Each of the elements in such a re-design exist in a complex causal network with the others, emphasising the importance of a coordinated and strategic rethink on the system-level. To reduce failure demand therefore leads to an abundance of meaningful projects that must create a synthetic whole that can address the underlying system weaknesses.

Of significance from a practical perspective, the failure demand that was identified in this thesis can be used as a justification for embarking on system improvement. In a way, the failure demand is a representation of how severe the consequences of the underlying problems are. This could act as a motivation to decision makers to justify the importance of reengineering the system, because the resources consumed because of failure demand are considerable. These resources include additional staff, warehouses that are larger than they need to be, more infrastructure, greater administrative complexity to track repeated orders, greater wear and tear on vehicles, higher cost to procure pharmaceuticals on the open market and a reduced quality of service and a variety of others. Failure demand can be treated as the bridge that ties the consumption of these resources to the underlying root causes.

7 - 4.2. Practical steps for implementation

Although implementation of improvements is beyond the scope of this doctoral work, it is important to provide commentary on how this could be done and what the difficulties are likely to be.

Any implementation of changes arising from this project would require a strategic system design. Failure demand may be created in instances in so-called "moments of truth" [44 Pg. 276] however the greater issue in this case is the system structure. This requires strategic interventions operating at the system planning – policy level; as described by Hines et al. [174]. For that reason, an intervention in this instance would need to focus on procurement strategies, government policy, and the overall supply-chain system.

In Section 2 - 6.3 of this thesis, a variety of sources were synthesised to present five key drivers for successful Lean implementations. Although these are not necessarily absolutes, they serve as a simple framework to understand potential challenges in implementation. These dimensions are "training", "quality data", "supportive culture", "involvement of managers" and "the buy-in of clinicians".

Training

To put "failure demand glasses" (a metaphor inspired by Ohno's "Muda glasses" [14]) on will require training and awareness building. It is proposed that suitable training programs are designed and that they are shared widely in organisations.

Green found that transferring knowledge to action is only successful in 8 % - 14 % of studied cases [431]. This is a low likelihood of success at the outset, however good training and the establishment of a knowledgeable team is a powerful way to improve the likelihood of success.

This is likely to be a multi-year project and will need a motivated skilled local team to implement the transition to a leaner organisation. Key transitions will need to be shared in training. The core ideas of Lean are intimidating to an organisation that is accustomed to hoarding and over-ordering. Learning to be more at ease with more regular smaller batches will require considerable knowledge transfer, mentorship and guidance. This will have to be introduced in conjunction with shorter time-frame forecasting techniques, and the improvement of consumption data.

Good reliable data

The ERP system has very good, reliable data, however this data can only reflect what was ordered and what was shipped. It does not reflect underlying demand patterns pulled from the patient. At all stages in the network: patients themselves, health facilities, secondary warehouses and main warehouses – hoarding takes place in anticipation of future stock-outs. For that reason, the comprehensive data is distorted by the underlying problem that true demand is misaligned with actual consumption.

Getting accurate data down to patient (or script) level is a daunting challenge. It is unreasonable to expect that this will be achieved in the short term, as the country lacks the necessary infrastructure (some health centres do not even have electricity or running water). Nevertheless, sampling studies can be used to establish rough baselines to start understanding the underlying demand patterns, which can be enriched with time, and tapping into the increasing access to mobile internet in the country.

This leaves two extremes, the comprehensive but exaggerated consumption data presented in the actual transit records and the incomplete but more accurate sampled data. Discretion and much work is required to find a middle ground that can serve as a basis for valid analysis to lead into a valid set of data about demand which the Vanguard method advocates as the departure point for improvement [32, 44].

A supportive culture

The agency responsible for pharmaceutical supply is highly hierarchically structured, with a strong command and control culture. Specialist roles and departments work in "back offices" with strict authoritarian processes, mistrustful management and a powerful bureaucracy. Decision making and authorisation is systematically the role of seniors and managers. The agency itself is a government department and as such itself also defers much of its commercial and logistic mandate to political leadership outside the agency.

The culture of the organisation represents a highly problematic impediment to implementation. Work will be required to adjust the culture and to move away from some of the command and control features that this system possesses [32], notably a strong top-down authoritarian approach to doing work.

The involvement of management

Involvement as well as overt support by management is important to make changes. This drives momentum of initiatives [111]. In this instance however, given that the initiative will be at first at a very high, strategic level, the involvement of management to encourage more junior staff to participate will be of reduced importance. As systemic changes start to emerge however, it will be important to engage directly with key gatekeeper managers. Resistance to apparently unfamiliar practices, such as placing smaller, less frequent, orders, will require strong management support. This is a major hurdle and will require considerable work to build up before any initiative commences.

Including clinicians

Clinicians operate at arms' length from this system and have limited power or influence in this system. However, the original sources [142, 163] justify the need to have clinicians involved in implementations, as clinicians typically act as gatekeepers, because they have the power to slow down or resist successful implementations by withholding their support, being mistrustful of results or prospective improvements. In this context the gate-keeping power resides with the political leadership outside of the agency. For that reason, it will be important to involve individuals at ministerial and permanent secretarial level.

7 - 5. General Impact of Failure Demand

7 - 5.1. Practical impact

On a practical level, this research has found, as expected, that failure demand leads to higher cost as all elements of the system need to be over-designed, over-staffed, and over-stocked to compensate for the inability of the system to satisfy the need placed on it.

Time is wasted in the system with processing and travel being unnecessarily required to compensate for the system's inability to satisfy the burden placed on it. Similarly, the system is over staffed, meaning that people are employed in roles to do work that would be unnecessary if resolution were reached on value demand. Questions around quality of care are difficult to answer [432], but this thesis proposes that due to failure demand, the ability of the system to provide a high quality service is compromised and that service quality could improve if failure demand were to drop. Similarly, on a practical level, in this case, it was found that failure demand leads to closed pharmacies, because pharmacists need to repeat journeys to return to buy pharmaceuticals that were unavailable at the previous instance of system failure. Pharmacists must drive considerable distances over difficult terrain to do so, which impacts on vehicle wear and tear, travel cost and eventual maintenance, breakdowns and the potential for accidents.

Based on these findings it is proposed that several transitions take place, in the delivery of service, anchored in policy, strengthened through training and recognising societal impacts.

7 - 5.2. Policy

To understand the policy implications of this work, Jacobson's idea of the "user group" [433, 434] is used. This concept proposes that the value of research can be expressed in terms of the consumers or users of the new knowledge. In this case, it is proposed that the user groups affected by this work crosses domains and includes users from frontline staff to policy makers. Although implementation is a site-level intervention, supporting change and putting in place failure demand limiting practices may require policy interventions.

This work has shown that reducing demand generally and failure demand, in particular, leads to a reduced burden on health facilities and a reduction in the need for capitalisation to address the gap between capacity and demand. this can lead to improved service delivery, more efficient operation and lower over-all cost.

To an extent, policy dictates practice [337]. Promoting systems thinking, with a focus on improvement, and particularly failure demand, into the policy mainstream, could encourage policy makers to foreground the importance hereof, and alter behaviours to match.

This work presents a case to argue for the prioritisation of improvement programs in health service delivery. It should do so by influencing the debate on service delivery [435] whilst guiding and shaping the practice for system design, including standardised approaches and philosophies, notably Lean and Systems Thinking. This will enhance the ability of decision makers to make evidence-based decisions.

Decision making based on an understanding of the demand structure leads to different design and implementation strategies to provide health care which leads to change in practice.

7 - 5.3. Society

This study only looks at the burden on the health system where deductions can be made. But there is a greater societal burden that cannot be ignored in the long run.

This was discussed in Section 5 - 7.3, page 221 but it is worth emphasising. Should a patient return, and become failure demand, the burden of the system failure is shifted to the health system. This thesis aims to limit this. However, other options remain, the patient may exit the system and then either heal without the formal medical system, stay the same or get worse.

In all cases other than failure demand, the burden of system failure shifts to society. This study found that a considerable amount of system failure is never translated into becoming failure demand. This raises several questions including the following thought experiment:

Does this mean that somewhere someone did not get their vital medications? Does this mean that somewhere someone died? Did a family lose a breadwinner? Is this thought experiment scalable? Does someone become many somebodies? Does many become hundreds, thousands? What is this burden? That question is difficult to assess and might require further investigations [16] – because that particular burden is recorded in societies, in families and in graveyards.

Although this study does not answer this question, it is an important one: What is the cost of failure demand on society?

7 - 6. Summary

Figure 7-31 (Enlarged in Appendix 6, Page 340) summarises the work done for this thesis. It structures the major contributions as they were formulated and summarises the overall impact and usefulness of this work. The image shows how the work was funnelled from the very broad: understanding the building blocks of demand in health care in general, to the fine, describing failure demand in a particular setting.

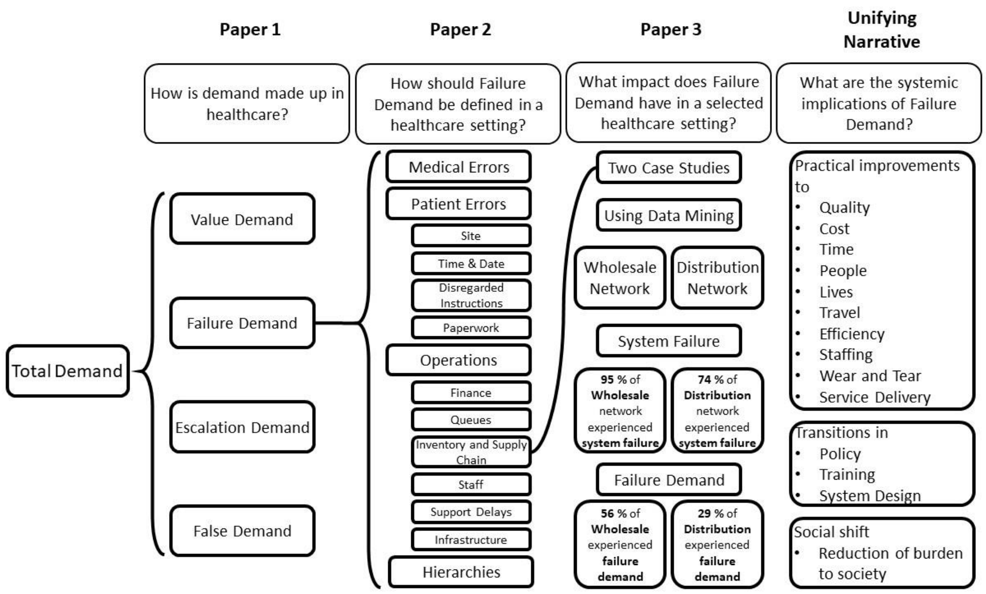


Figure 7-31: Summary of the major contributions made in this doctorate

Demand is complex, it consists of many different modalities. A population model is introduced that consists of an unhealthy and a healthy population, the demand population within this framework is defined. Out of this, demand is granularized into four categories, value demand, failure demand, escalation demand and false demand. Failure demand is chosen for specific elaboration. A mental model is created that classifies failure modes as system failure or not – and then using one of these failure modes – drug shortages, a double case study is performed that explores the health system in a developing country. It was discovered that the burden placed on the pharmaceutical supply-chain in the selected country is considerable and the impact of system failure is a specific and significant contributor to the fragility of the provision of service to vulnerable communities in this case.

7 - 6.1. Limitations of this study

Failure demand is difficult to measure and requires considerable amounts of complete and comprehensive data. Although large patient data sets were made available for the purposes of this study, these were not used, as the data integrity was too poor to make defendable deductions. As a result, this study does not present findings on failure demand in a clinical setting. This limits the clinical assessment of failure demand and hence the extent to which results obtained in the study are transferable to clinical settings which means confidence of findings is reduced for clinical settings and this aspect remains but of future interest. Repeating similar work in a clinical setting tracking the failure demand of patients is an opportunity to make a further contribution. Although clinical assessment was abandoned due to poor data, the software was developed that would be able to find failure demand in more reliable clinical data sets.

Furthermore, due to restrictions imposed by the Human Research Ethics Committee which resists granting access to patients for non-clinically trained researchers, this study is purposefully scoped to exclude the need to interact directly with patients. This limits the ability to enrich findings making use of rich observational primary data to improve reliability and validity of the work through a triangulated approach.

This work is conducted in a general health care setting, considering contexts from experience and from literature. The vast diversity of health care contexts globally are not represented in these sources work and as such behaviours beyond observation and literature are likely to exist. Experience and observation can further be limiting by definition. For this reason, caution must be taken and these results must not be presumed to apply to all similar contexts.

The two empirical case studies in a single developing country cannot be generalised. The findings from this study should be considered cautiously and not interpreted to represent other markets or other countries, even those with similar conditions.

This thesis is limited by context. Although the mental models presented are intended to be generally applicable, they are biased towards the context of familiarity of the author, which is primarily in six developing countries' health systems. Health care systems can differ fundamentally so some of the models' categories may require modification by future researchers for their national and sectoral research contexts.

This study examined the systems questions related to health service delivery improvement. The author is not clinically trained, and hence all clinical matters are out of scope.

Failure demand is a significant topic and as identified in Chapter 5 can have a genesis from a diversity of origins. The pharmaceutical supply-chain offered a good opportunity to measure failure demand in an environment with a significant amount of reliable data, however outcomes from one origin may not apply to another and so caution in interpretation needs to be exercised.

Although the poor supply-chain-management is presumed to be a part of the system, complex, self-organising systems with emergent characteristics means that agents external to the formal supply-chain system may play a role, including vendors, political interests, international organisations, and a diversity of other actors that are not considered.

7 - 7. Contributions.

7 - 7.1. Major contributions

This thesis has contributed to knowledge in several ways. The major contributions emerging from this research are

2. The introduction of a framework that explores demand from the point of view of systemic mechanisms, leading to a demand modularisation model

3. Shifting the focus to system failures, rather than human failures or motives as the exacerbators of demand and the seed for failure demand,

4. Introducing the idea of escalation demand a previously un-described demand modality stemming from the increasing demand due to delay,

5. Seddon's thinking is further enriched by adding a hierarchical demand modality which is an undescribed mode of failure demand,

6. Producing an identification framework and a list of common practices that lead to system failure,

7. Using these contributions to provide an understanding of what impact failure demand has on a selected pharmaceutical supply-network in a developing country,

8. Establishing benchmark failure demand values in a health care setting, a domain with little reported information on failure demand.

7 - 7.2. Other contributions

Additional contributions were also made, without which the research framing for the major contribution could not have been made. These include:

1. A definition for what an unhealthy person is, defined as the natural set that should be the feeder for demand in health care,

2. A mental model that defines the social dynamic between healthy and unhealthy sets of the population and defines the originating population for health system demand from these,

3. Creating two data mining tools that may be fully transferrable, or otherwise with minor modification adaptable to any similar environment to assess failure demand.

7 - 8. Conclusion

The main objectives of this study were to:

1. Create a demand modality model based on the systemic mechanisms that are the root causes for demand,
2. Define failure demand for a complex hierarchical system, with consideration of the health system,
3. Measure an example of the prevalence of failure demand in a health setting.

These objectives have been achieved through the following chapters:

In Chapter 4, this thesis presents a demand modularisation that improves Vickery & Lynch' [19] demand model by overlaying Seddon's modality of failure demand, however the thesis enriches this further through the introduction of escalation demand, a mode foreign to most other service environments. By doing so, the perspective that considers the mechanisms of root causes leading to demand is introduced.

In Chapter 5, a model derived from literature and experience describes the nature of failure demand in health systems. The importance of system failure is discussed, and the triggers are described. The causal relationship between system failures as the root cause and failure demand as the symptom [436] is emphasised. This model introduces a structured hierarchical definition for system failure in health care, introduces the idea of failure demand II, the mode of failure demand that crosses hierarchical bridges. This enables an overall systems-view of health care facilities including interconnectedness among facilities, which has policy implications.

Failure demand, in the literature, has been found to range between 40 % and 80 % of total demand [44]. In Chapter 6 a study of a national pharmaceutical supply-chain, that made use of data mining techniques showed that failure demand does account for a considerable part of the system burden. This opens the conversation on failure demand in pharmaceutical supply-chains.

Regardless of the quantum of failure demand, systematic interventions can reduce demand, meaning that capitalisation and staffing could be reduced, or more usefully, service levels could improve at no additional cost. This, in the face of considerable cost pressures on health systems globally, is a desirable pursuit.

Although of theoretical interest, understanding failure demand in health systems is potentially of tremendous practical value also. Given that reduced failure demand could have a dramatic impact on demand patterns, it is recommended that health system strengthening policy incorporates the reduction of system failures.

As noted earlier, the work conducted in this study has achieved the objectives of the research. Yet much work remains to strengthen the arguments about failure demand in health systems.

7 - 9. Recommendations

The work proposed here suggests that health care can be engineered as a system. That if the same attention is paid to managing demand, as is paid to managing capacity, considerable systemic improvements could be made.

Because failure demand is one of the major contributors to health care system load, it is proposed that if failure demand is reduced, lives could be improved.

If, however no changes are made, the status quo remains, which includes higher demand than necessary, the unnecessary hiring of more staff, the unnecessary building of more health facilities, the unnecessary buying of more infrastructure, essentially, unnecessary spending, on a grand scale.

It is recommended that further work is done to:

1. Assess the prevalence of failure demand in health systems by:
    a. Assessing all the modalities of failure demand in separate studies,
    b. Repeating this study in other settings, including public and private institutions and also in other developing and developed countries.
2. Study the impact of Escalation demand to gauge its cost and burden on a health system,
3. Lobby hospitals and health systems to keep stringent records that keep track of patient IDs, diagnostic codes etc. allowing for rigorous and otherwise improved analysis,
4. Make use of the data mining tool that was developed for clinical admissions records on data sets that have adequate integrity to establish true benchmark values that are not distorted by underlying data poverty,

5. Perform a failure demand study using the definitions and theoretical frameworks from this thesis in a clinical setting. This is subject to obtaining more complete health records,
6. Developing, circulating and training failure demand training materials as part of health systems strengthening thinking.

The efficient service system must limit failure demand. It is like a manufacturing system limiting rework.

Failure demand should be just as visible as parts in a reject bin, failure demand is thick waiting list books, overflowing waiting areas, long queues, tired doctors and rings under nurses' eyes. To say that failure demand is not visible, is to say that we do not see these things. ~ DH

---

7 - 10. References

14. Ohno, T., *Toyota Production System: Beyond Large-Scale Production*. 1988, CRC Press, Boca Raton, London, New York
15. Liker, J.K., *The Toyota Way*. 2004, McGraw Hill, New York
16. Yin, R.K., *Case Study Research: Design and Methods*. 1987, Sage Publications, Inc., Beverley Hills
19. Vickery, D.M. and Lynch, W.D., *Demand Management: Enabling Patients to Use Medical Care Appropriately.* Journal of Occupational and Environmental medicine, 1995. 37(5), p. 551-557.
30. Moraros, J., Lemstra, M., and Nwankwo, C., *Lean Interventions in Healthcare: Do They Actually Work? A Systematic Literature Review.* International Journal for Quality in Health Care, 2016. 28(2), p. 150-165.
32. Seddon, J., *The Vanguard Method.* Introduction to Vanguard - eLearning, 2020(1).
44. Seddon, J., *Freedom from Command and Control: Rethinking Management for Lean Service*. 2005, CRC press, Boca Raton, Florida
63. Bicheno, J. and Holweg, M., *The Lean Toolbox: A Handbook for Lean Transformations*. 5th Edition ed. 2016, PICSIE Books, Buckingham, UK
65. Womack, J.P., Jones, D.T., and Roos, D., *The Machine That Changed the World: How Lean Production Revolutionized the Global Car Wars*. 1991, Simon & Schuster, London
69. Seddon, J., *Watch out for the Toolheads.* Lean-service. com [accessed 1 Feb 2005], 2005.
92. Williams, A., Kennedy, S., Philipp, F., and Whiteman, G., *Systems Thinking: A Review of Sustainability Management Research.* Journal of Cleaner Production, 2017. 148, p. 866-881.
96. Checkland, P., *Systems Thinking.* Rethinking Management Information Systems, 1999, p. 45-56.
97. Waltner-Toews, D., Kay, J., Kay, J.J., and Lister, N.-M.E., *The Ecosystem Approach: Complexity, Uncertainty, and Managing for Sustainability*. 2008, Columbia University Press
99. Senge, P.M., *The Fifth Discipline.* Measuring Business Excellence, 1997. 1(3), p. 46-51.
101. Ackoff, R.L., Ackoff, R.L., and Emery, F.E., *On Purposeful Systems: An Interdisciplinary Analysis of Individual and Social Behavior as a System of Purposeful Events*. 2005, Transaction Publishers

105. Rusoja, E., Haynie, D., Sievers, J., Mustafee, N., Nelson, F., Reynolds, M., et al., *Thinking About Complexity in Health: A Systematic Review of the Key Systems Thinking and Complexity Ideas in Health.* Journal of Evaluation in Clinical Practice, 2018. 24(3), p. 600-606.
110. Piercy, N. and Rich, N., *Lean Transformation in the Pure Service Environment: The Case of the Call Service Centre.* International Journal of Operations & Production Management, 2009. 29(1), p. 54-76.
111. Gupta, S. and Sharma, M., *Lean Services: A Systematic Review.* International Journal of productivity and performance management, 2016.
119. Seddon, J., *Watch out for the Toolheads! Everything You Need to Know About Lean Manufacturing Tools and Why They Won't Work in Service Organisations*. 2005, Systemthinkingmethod.com.
138. Seddon, J., *Lean Is a Waning Fad.* Management Services, 2011. 55(4), p. 34-36.
142. Antony, J., Sreedharan, R., Chakraborty, A., and Gunasekaran, A., *A Systematic Review of Lean in Healthcare: A Global Prospective.* International Journal of Quality & Reliability Management, 2019.
163. Andersen, H., Røvik, K.A., and Ingebrigtsen, T., *Lean Thinking in Hospitals: Is There a Cure for the Absence of Evidence? A Systematic Review of Reviews.* BMJ open, 2014. 4(1).
174. Hines, P., Holweg, M., and Rich, N., *Learning to Evolve: A Review of Contemporary Lean Thinking.* International Journal of Operations & Production Management, 2004. 24(10), p. 994-1011.
184. Seddon, J., *Discussion of Failure Demand Repeat Consumption and the Idea of Demand*, D. Hartmann, Editor. 2020, Johannesburg, South Africa.
185. Marr, B. and Neely, A., *Managing and Measuring for Value: The Case of Call Centre Performance*. 2004, Cranfield School of Management and Fujitsu, Cranfield, UK.
186. Walley, P., *Does the Public Sector Need a More Demand-Driven Approach to Capacity Management?* Production Planning & Control, 2013. 24(10-11), p. 877-890.
192. Allder, S., Walley, P., and Silvester, K., *Is Follow-up Capacity the Current Nhs Bottleneck?* Clinical Medicine, 2011. 11(1), p. 31.

194. Walley, P., Found, P., and Williams, S., *Failure Demand: A Concept Evaluation in Uk Primary Care.* International Journal of Health Care Quality Assurance, 2019. 32(1), p. 21-33.
210. Jack, E.P. and Powers, T.L., *A Review and Synthesis of Demand Management, Capacity Management and Performance in Health-Care Services.* International Journal of Management Reviews, 2009. 11(2), p. 149-174.
235. Seddon, J. and Brand, C., *Debate: Systems Thinking and Public Sector Performance.* 2008.
239. Deming, E., *Out of the Crisis. Quality, Productivity and Competitive Position. Cambridge, Ma: Center for Advanced Engineering Study*. 1986, Cambridge University Press.
240. Masters, K.I.A., *From New Public Management to Lean Thinking: Understanding and Managing'potentially Avoidable Failure Induced Demand'*. 2010, University of Sussex.
241. Lillrank, P., Groop, P.J., and Malmström, T.J., *Demand and Supply–Based Operating Modes—a Framework for Analyzing Health Care Service Production.* Milbank Quarterly, 2010. 88(4), p. 595-615.
320. Seddon, J. and Caulkin, S., *Systems Thinking, Lean Production and Action Learning.* Action Learning: Research and Practice, 2007. 4(1), p. 9-24.
337. Bowen, S. and Zwi, A.B., *Pathways to "Evidence-Informed" Policy and Practice: A Framework for Action.* PLoS Med, 2005. 2(7).
425. Rother, M. and Shook, J., *Learning to See: Value Stream Mapping to Add Value and Eliminate Muda*. 2003, Lean Enterprise Institute
426. Richardson, T. and Richardson, E., *The Toyota Engagement Equation: How to Understand and Implement Continuous Improvement Thinking in Any Organization*. 1st edition ed. 2017, McGraw-Hill Education, New York City
427. World Health Organization, *Everybody's Business--Strengthening Health Systems to Improve Health Outcomes: Who's Framework for Action.* 2007.
428. Lansing, J.S., *Complex Adaptive Systems.* Annual Review of Anthropology, 2003. 32(1), p. 183-204.
429. Agrawal, R. and Srikant, R. *Mining Sequential Patterns*. in *International Conference on Data Engineering*. 1995.
430. Mabroukeh, N.R. and Ezeife, C.I., *A Taxonomy of Sequential Pattern Mining Algorithms.* ACM Computing Surveys (CSUR), 2010. 43(1), p. 3.

431. Green, L.W., *Public Health Asks of Systems Science: To Advance Our Evidence-Based Practice, Can You Help Us Get More Practice-Based Evidence?* American Journal of Public Health, 2006. 96(3), p. 406-409.
432. Campbell, S.M., Roland, M.O., and Buetow, S.A., *Defining Quality of Care.* Social Science & Medicine, 2000. 51(11), p. 1611-1625.
433. Jun, J.B., Jacobson, S.H., and Swisher, J.R., *Application of Discrete-Event Simulation in Health Care Clinics: A Survey.* The journal of the Operational Research Society, 199. 50(2), p. 109-123.
434. Jacobson, N., Butterill, D., and Goering, P., *Development of a Framework for Knowledge Translation: Understanding User Context*. Journal of Health Services Research & Policy, 2003. 8(2), p. 94-99.
435. Reid, P.P., Compton, W.D., Grossman, J.H., and Fanjiang, G., *Building a Better Delivery System: A New Engineering/Health Care Partnership*. 2005, National Academies Press, Washington D.C.
436. Andersen, B. and Fagerhaug, T., *Root Cause Analysis: Simplified Tools and Techniques*. 2006, ASQ Quality Press, Milwaukee, Wisconsin

Chapter Eight: **Full Reference List for the Thesis**

1. Merry, M.D., *Healthcare's Need for Revolutionary Change.* Quality Progress, 2003. 36(9), p. 31-35.
2. World Health Organization, *World Health Statistics 2017*. 2017, World Health Organization
3. Worldometer. *Government and Economics Indicators*. 2020 [cited 2020 15/03/2020].
4. International Monetary Fund, *World Economic Outlook*. 2020, International Monetary Fund, Washington D.C.
5. United Nations Development Programme, *Human Development Report 2019*. 2019, United Nations, New York City - New York,
6. Unesco Institute for Statistics, *Global Education Digest*. 2020, UIS Montreal.
7. Sarnak, D.O., Squires, D., Kuzmak, G., and Bishop, S., *Paying for Prescription Drugs around the World: Why Is the Us an Outlier.* Commonwealth Fund Issue Briefs, 2017. 2017, p. 1-14.
8. Dijulio, B., Kirzinger, A., Wu, B., and Brodie, M., *Data Note: Americans' Challenges with Health Care Costs.* Kaiser Family Foundation, 2017.
9. Whitehurst, L., *Utah Sends Employees to Mexico for Lower Prescription Prices*, in *Associated Press News*. 2020, The Associated Press, New York.
10. Affordable Care Act, *Public Law 111-148*, in *Title IV, x4207, USC HR*. 2010, United States of America. p. 2010.
11. Smith, M., Saunders, R., Stuckhardt, L., and Mcginnis, J.M., *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*. 2013, National Academies Press, Washington D.C.
12. Aaron, H.J. and Schwartz, W.B., *Managed Competition: Little Cost Containment without Budget Limits.* Health Affairs, 1993. 12(1), p. 204-215.
13. Graban, M., *Lean Hospitals*. 2008, CRC Press, Boca Raton, London, New York
14. Ohno, T., *Toyota Production System: Beyond Large-Scale Production*. 1988, CRC Press, Boca Raton, London, New York
15. Liker, J.K., *The Toyota Way*. 2004, McGraw Hill, New York
16. Yin, R.K., *Case Study Research: Design and Methods*. 1987, Sage Publications, Inc., Beverley Hills
17. Levine, H.G., Gallimore, R., Weisner, T.S., and Turner, J.L., *Teaching Participant-Observation Research Methods: A Skills-Building*

Approach. Anthropology & Education Quarterly, 1980. 11(1), p. 38-54.
18. Moodie, J. and Hapgood, K. *The Phd Thesis by Publication in Engineering: Insights for Supervisors*. in *23rd Annual Conference of the Australasian Association for Engineering Education 2012: Profession of Engineering Education: Advancing Teaching, Research and Careers, The*. 2012. Engineers Australia.
19. Vickery, D.M. and Lynch, W.D., *Demand Management: Enabling Patients to Use Medical Care Appropriately.* Journal of Occupational and Environmental medicine, 1995. 37(5), p. 551-557.
20. Marczyk, G., Dematteo, D., and Festinger, D., *Essentials of Research Design and Methodology*. 2005, John Wiley & Sons Inc, Hoboken, New Jersey
21. Beall, J. *Beall's List* Beall's List of Predatory Journals and Publishers 2020 [cited 2020.
22. Isi Web of Science, *Isi Web of Science Database*. 2020, Clarivate Analytics, Philadelphia, USA.
23. Adams, R.J., Smart, P., and Huff, A.S., *Shades of Grey: Guidelines for Working with the Grey Literature in Systematic Reviews for Management and Organizational Studies.* International Journal of Management Reviews, 2017. 19(4), p. 432-454.
24. Conn, V.S., Valentine, J.C., Cooper, H.M., and Rantz, M.J., *Grey Literature in Meta-Analyses.* Nursing Research, 2003. 52(4), p. 256-261.
25. Mcauley, L., Tugwell, P., and Moher, D., *Does the Inclusion of Grey Literature Influence Estimates of Intervention Effectiveness Reported in Meta-Analyses?* The Lancet, 2000. 356(9237), p. 1228-1231.
26. Dickersin, K., Min, Y.-I., and Meinert, C.L., *Factors Influencing Publication of Research Results: Follow-up of Applications Submitted to Two Institutional Review Boards.* JAMA, 1992. 267(3), p. 374-378.
27. Dickersin, K. and Min, Y.-I., *Nih Clinical Trials and Publication Bias.* The Online Journal of Current Clinical Trials, 1993.
28. Easterbrook, P.J., Gopalan, R., Berlin, J., and Matthews, D.R., *Publication Bias in Clinical Research.* The Lancet, 1991. 337(8746), p. 867-872.
29. Rotonda, T., *Underreporting Research Is Scientific Misconduct.* JAMA, 1990. 263, p. 1405-1408.

30. Moraros, J., Lemstra, M., and Nwankwo, C., *Lean Interventions in Healthcare: Do They Actually Work? A Systematic Literature Review.* International Journal for Quality in Health Care, 2016. 28(2), p. 150-165.
31. Crotty, M., *The Foundations of Social Research: Meaning and Perspective in the Research Process*. 1998, Sage, New York City
32. Seddon, J., *The Vanguard Method.* Introduction to Vanguard - eLearning, 2020(1).
33. Seddon, J., *Failure Demand.* 2020.
34. Kingman, J.F.C., *The Single Server Queue in Heavy Traffic.* Mathematical Proceedings of the Cambridge Philosophical Society, 1961. 57(04), p. 902-904.
35. Kingman, J.F., *Two Similar Queues in Parallel.* The Annals of Mathematical Statistics, 1961. 32(4), p. 1314-1323.
36. Kingman, J.F.C., *On Queues in Heavy Traffic.* Journal of the Royal Statistical Society. Series B (Methodological), 1962. 24(2), p. 383-392.
37. Kingman, J.F.C., *On the Algebra of Queues.* Journal of Applied Probability, 1966. 3(2), p. 285-326.
38. Bicheno, J. *Towards Reducing Queues: Muri, Mura, Muda*. in *European Lean Educator Conference (ELEC2018)*. 2018. Braga, Portugal.
39. Bicheno, J., *Queues!*, D. Hartmann, Editor. 2019, Buckingham, United Kingdom.
40. Hopp, W.J. and Spearman, M.L., *Factory Physics*. 2011, Waveland Press, Long Grove, Illinois
41. Bicheno, J., *The King of Equations*, in *The Lean Management Journal*. 2011, SayOne Media, Norwich. p. 6.
42. Bicheno, J., *The Lean Games and Simulations Book*. 2015, Picsie Books, Buckingham, UK
43. Roser, C., *A Small Dice Game for the Kingman Formula*, in *All About Lean*, C. Roser, Editor. 2017.
44. Seddon, J., *Freedom from Command and Control: Rethinking Management for Lean Service*. 2005, CRC press, Boca Raton, Florida
45. Little, J.D. and Graves, S.C., *Little's Law*, in *Building Intuition*. 2008, Springer, New York City. p. 81-100.
46. Morse, P.M., *Queues, Inventories, and Maintenance*. 1958, Dover Publications Inc., Mineola, New York

47. Little, J.D., *A Proof for the Queuing Formula.* Operations Research, 1961. 9(3), p. 383-387.
48. Potter, A., Towill, D.R., and Gosling, J., *On the Versatility of Little's Law in Operations Management: A Review and Classification Using Vignettes.* Production Planning & Control, 2019, p. 1-16.
49. Jones, S.S., Thomas, A., Evans, R.S., Welch, S.J., Haug, P.J., and Snow, G.L., *Forecasting Daily Patient Volumes in the Emergency Department.* Academic Emergency Medicine, 2008. 15(2), p. 159-170.
50. Goldratt, E.M., *The Goal: A Process of Ongoing Improvement,* . 2004, Routledge, New York City
51. Brougham, G., *The Cynefin Mini-Book*. 2015, Lulu. com, Morrisville, North Carolina
52. Kurtz, C.F. and Snowden, D.J., *The New Dynamics of Strategy: Sense-Making in a Complex and Complicated World.* IBM Systems Journal, 2003. 42(3), p. 462-483.
53. Snowden, D., *Strategy in the Context of Uncertainty.* Handbook of Business strategy, 2005. 6(1), p. 47-54.
54. Snowden, D.J. and Boone, M.E., *A Leader's Framework for Decision Making.* Harvard Business Review, 2007. 85(11), p. 68.
55. Hasan, H. and Kazlauskas, A., *Making Sense of Is with the Cynefin Framework.* University of Wolongong, 2009.
56. Geron, S.M., *21st Century Strategies for Policing Protest: What Major Cities' Responses to the Occupy Movement Tell Us About the Future of Police Response to Public Protest*. 2014, Naval Postgraduate School, Monterey, California.
57. Snowden, D., *Journal of Knowledge Management.* Journal of knowledge management, 2002. 6(2), p. 100-111.
58. Lowy, A. and Hood, P., *The Power of the 2 X 2 Matrix: Using 2 X 2 Thinking to Solve Business Problems and Make Better Decisions*. 2011, John Wiley & Sons, Hoboken, New Jersey
59. Madsen, D.O., *Not Dead Yet: The Rise, Fall and Persistence of the Bcg Matrix.* Problems and Perspectives in Management, 2017. 15(1), p. 19-34.
60. Vollmar, J., Gepp, M., Palm, H., and Cala, A. *Engineering Framework for the Future: Cynefin for Engineers*. in *2017 IEEE International Systems Engineering Symposium (ISSE)*. 2017. IEEE.
61. Hasan, H. and Kazlauskas, A., *The Cynefin Framework: Putting Complexity into Perspective.* Univerity of Wolongong, 2014.

62. Snowden, D., *The Knowledge You Need, Right When You Need It.* Knowledge Management Review, 2003. 5, p. 24-27.
63. Bicheno, J. and Holweg, M., *The Lean Toolbox: A Handbook for Lean Transformations*. 5th Edition ed. 2016, PICSIE Books, Buckingham, UK
64. Jones, D.T. and Womack, J.P., *Lean Thinking*. 2012, Simon & Shuster, London
65. Womack, J.P., Jones, D.T., and Roos, D., *The Machine That Changed the World: How Lean Production Revolutionized the Global Car Wars*. 1991, Simon & Schuster, London
66. Krafcik, J.F., *Triumph of the Lean Production System.* MIT Sloan Management Review, 1988. 30(1), p. 41.
67. Lean Enterprise Institute. *A Brief History of Lean*. 2019 2019/04/02]; Available from: https://www.lean.org/WhatsLean/History.cfm.
68. Dolinsky, A., *Inventory Management History Part Three: Venetian Arsenal–Ahead of Their Time,*. 2013, Almyta Systems, Reno, Nevada.
69. Seddon, J., *Watch out for the Toolheads.* Lean-service. com [accessed 1 Feb 2005], 2005.
70. Smith, A. and Nicholson, J.S., *An Inquiry into the Nature and Causes of the Wealth of Nations*. 1887, T. Nelson and Sons, Edinburgh, UK
71. Whitney, E., *Correspondence of Eli Whitney Relative to the Invention of the Cotton Gin*. 2013, Read Books Ltd, Worcestershire, UK
72. Mccarthy, M. and Garcia, R., *Screw Threads on the Ss Xantho Engine: A Case of Standardisation in 19th Century Britain.* International Journal of Nautical Archaeology, 2004. 33(2), p. 330-337.
73. Woodbury, R.S., *The Legend of Eli Whitney and Interchangeable Parts.* Technology and Culture, 1960. 1(3), p. 235-253.
74. Gilbreth, F.B., *Motion Study: A Method for Increasing the Efficiency of the Workman*. 1911, D. Van Nostrand Company, New York City
75. Wren, D.A. and Bedeian, A.G., *The Evolution of Management Thought*. 1994, John Wiley & Sons, Hoboken, New Jersey
76. Fayol, H., *General and Industrial Management*. 2016, Ravenio Books, Cambridge, UK
77. Gilbreth, F.B., *Primer of Scientific Management*. 1912, D. Van Nostrand Co., New York
78. Taylor, F.W., *The Principles of Scientific Management*. 1914, Harper, New York

79. Gilbreth, F.B., *Bricklaying System*. 1909, The M.C. Clark publishing co., New York City
80. Gilbreth, F.B. and Gilbreth, L.M., *Fatigue Study, the Elimination of Humanity's Greatest Unnecessary Waste: A First Step in Motion Study*. 1916, Sturgis & Walton Company, New York City
81. Locke, E.A., *The Ideas of Frederick W. Taylor: An Evaluation.* Academy of Management Review, 1982. 7(1), p. 14-24.
82. Gilbreth, F.B. and Gilbreth, L.M., *Applied Motion Study a Collection of Papers on the Efficient Method to Industrial Preparedness.* 1917.
83. Khamara, E.J., *Space, Time, and Theology in the Leibniz-Newton Controversy*. Vol. 6. 2013, Walter de Gruyter
84. Cassirer, E., *Newton and Leibniz.* The Philosophical Review, 1943. 52(4), p. 366-391.
85. Taylor, F.W., *Scientific Management*. 2004, Routledge, Abingdon, UK
86. Ford, H., *Today and Tomorrow: Commemorative Edition of Ford's 1926 Classic*. 2018, Routledge, Abingdon, UK
87. Levinson, W.A., *Henry Ford's Lean Vision: Enduring Principles from the First Ford Motor Plant*. 2012, Productivity Press, New York City
88. Arnheiter, E.D. and Maleyeff, J., *The Integration of Lean Management and Six Sigma.* The TQM Magazine, 2005.
89. Mangaroo-Pillay, M., *A South African Ubuntu Analogy of the Lean Philosophy*, in *Department of Industrial Engineering, The North West University*. 2020, The North West University, Potchefstroom, South Africa.
90. Womack, J.P. and Jones, D.T., *Lean Thinking: Banish the Waste and Create in Your Corporation*. 1996, Simon and Schuster, London
91. Zidel, T.G., *A Lean Guide to Transforming Healthcare*. 2006, ASQ Quality Press, Milwaukee, Wisconsin
92. Williams, A., Kennedy, S., Philipp, F., and Whiteman, G., *Systems Thinking: A Review of Sustainability Management Research.* Journal of Cleaner Production, 2017. 148, p. 866-881.
93. Smuts, J.C., *Holism and Evolution.* 1926.
94. Becht, G., *Systems Theory, the Key to Holism and Reductionism.* Bioscience, 1974. 24(10), p. 569-579.
95. Checkland, P., *Systems Theory and Management Thinking.* American Behavioral Scientist, 1994. 38(1), p. 75-91.
96. Checkland, P., *Systems Thinking.* Rethinking Management Information Systems, 1999, p. 45-56.

97. Waltner-Toews, D., Kay, J., Kay, J.J., and Lister, N.-M.E., *The Ecosystem Approach: Complexity, Uncertainty, and Managing for Sustainability*. 2008, Columbia University Press
98. Bicheno, J., *The Service Systems Toolbox: Integrating Lean Thinking, Systems Thinking and Design Thinking*. 2012, Picsie Books, Buckingham, UK
99. Senge, P.M., *The Fifth Discipline.* Measuring Business Excellence, 1997. 1(3), p. 46-51.
100. Ackoff, R.L., *Management Misinformation Systems.* Management Science, 1967. 14(4), p. B-147-B-156.
101. Ackoff, R.L., Ackoff, R.L., and Emery, F.E., *On Purposeful Systems: An Interdisciplinary Analysis of Individual and Social Behavior as a System of Purposeful Events*. 2005, Transaction Publishers
102. Checkland, P., *From Optimizing to Learning: A Development of Systems Thinking for the 1990s.* The journal of the Operational Research Society, 1985. 36(9), p. 757-767.
103. Nakagawa, C.R.C., Nordholm, S., and Yan, W.-Y., *Analysis of Two Microphone Method for Feedback Cancellation.* IEEE Signal Processing Letters, 2014. 22(1), p. 35-39.
104. Homer, J.B. and Hirsch, G.B., *System Dynamics Modeling for Public Health: Background and Opportunities.* American Journal of Public Health, 2006. 96(3), p. 452-458.
105. Rusoja, E., Haynie, D., Sievers, J., Mustafee, N., Nelson, F., Reynolds, M., et al., *Thinking About Complexity in Health: A Systematic Review of the Key Systems Thinking and Complexity Ideas in Health.* Journal of Evaluation in Clinical Practice, 2018. 24(3), p. 600-606.
106. Lorenz, E., *The Butterfly Effect.* World Scientific Series on Nonlinear Science Series A, 2000. 39, p. 91-94.
107. Best, A. and Holmes, B., *Systems Thinking, Knowledge and Action: Towards Better Models and Methods.* Evidence & Policy: A Journal of Research, Debate and Practice, 2010. 6(2), p. 145-159.
108. Betzler, S.B., Wisnet, A., Breitbach, B., Mitterbauer, C., Weickert, J., Schmidt-Mende, L., et al., *Template-Free Synthesis of Novel, Highly-Ordered 3d Hierarchical Nb 3 O 7 (Oh) Superstructures with Semiconductive and Photoactive Properties.* Journal of Materials Chemistry A, 2014. 2(30), p. 12005-12013.
109. Luhmann, N., *Social Systems*. 1995, stanford university Press

110. Piercy, N. and Rich, N., *Lean Transformation in the Pure Service Environment: The Case of the Call Service Centre.* International Journal of Operations & Production Management, 2009. 29(1), p. 54-76.
111. Gupta, S. and Sharma, M., *Lean Services: A Systematic Review.* International Journal of productivity and performance management, 2016.
112. Sasser, W.E., *Match Supply and Demand in Service Industries.* Harvard Business Review, 1976. 54(6), p. 133-140.
113. Levitt, T., *Production-Line Approach to Service.* Harvard Business Review, 1972. 50(5), p. 41-52.
114. Levitt, T., *Industrialization of Service.* Harvard Business Review, 1976. 54(5), p. 63-74.
115. Radnor, Z. and Walley, P., *Learning to Walk before We Try to Run: Adapting Lean for the Public Sector.* Public Money and Management, 2008. 28(1), p. 13-20.
116. Radnor, Z. and Osborne, S.P., *Lean: A Failed Theory for Public Services?* Public Management Review, 2013. 15(2), p. 265-287.
117. Bowen, D.E. and Youngdahl, W.E., *"Lean" Service: In Defense of a Production-Line Approach.* International Journal of Service Industry Management, 1998.
118. Spear, S.J., *Fixing Health Care from the inside, Today.* Harvard Business Review, 2005. 83(9), p. 78.
119. Seddon, J., *Watch out for the Toolheads! Everything You Need to Know About Lean Manufacturing Tools and Why They Won't Work in Service Organisations*. 2005, Systemthinkingmethod.com.
120. Radnor, Z., Walley, P., Stephens, A., and Bucci, G., *Evaluation of the Lean Approach to Business Management and Its Use in the Public Sector.* Scottish Executive Social Research, 2006. 20, p. 1-6.
121. Jones, D.T. and Womack, J.P., *Lean Consumption.* Harvard Business Review, 2005. 83, p. 58-68.
122. Womack, J., *Lean Consumption.* Manufacturing Engineer, 2005. 84(4), p. 5-5.
123. Allway, M. and Corbett, S., *Shifting to Lean Service: Stealing a Page from Manufacturers' Playbooks.* Journal of Organizational Excellence, 2002. 21(2), p. 45-54.
124. Oed Online, *The Oxford English Dictionary*, in *The Oxford English Dictionary*. 2020, Oxford University Press, Oxford, UK.

125. Cobra, M., *Serviços: Como Contruir Valor Para O Cliente*. 2004, Marcos Cobra Editora Ltda
126. Leite, H.D.R. and Vieira, G.E., *Lean Philosophy and Its Applications in the Service Industry: A Review of the Current Knowledge.* Production, 2015. 25(3), p. 529-541.
127. George, M., *Lean Six Sigma for Service*. 2003, McGraw Hill Professional, New York City
128. Johnston, R., *Service Operations Management: Return to Roots.* International Journal of Operations & Production Management, 1999. 19(2), p. 104 - 124.
129. Andrés-López, E., González-Requena, I., and Sanz-Lobera, A., *Lean Service: Reassessment of Lean Manufacturing for Service Activities.* Procedia Engineering, 2015. 132(1), p. 23-30.
130. Roy, R.B., Lillrank, P., Sreekanth, V., and Torkki, P., *The Conceptual Tools*, in *Designing Service Machines*. 2019, Springer. p. 5-19.
131. Roy, R.K., *A Primer on the Taguchi Method*. 2010, Society of Manufacturing Engineers, Ontario, Canada
132. Teece, D.J., *Business Models, Business Strategy and Innovation.* Long Range Planning, 2010. 43(2-3), p. 172-194.
133. Schneider, D. and Schnetkamp, G., *E-Markets: B2b-Strategien Im Electronic Commerce, Marktplätze, Fachportale, Plattformen*. 2013, Springer-Verlag, Wiesbaden
134. Seddon, J. and Brand, C., *Debate: Systems Thinking and Public Sector Performance.* Public Money & Management, 2008. 28(1), p. 7-9.
135. Skinner, W., *Manufacturing-Missing Link in Corporate Strategy.* Harvard Business Review, 1969, p. 136 - 145.
136. Seddon, J., *I Want You to Cheat!: The Unreasonable Guide to Service and Quality in Organisation*. 1 ed. 1992, Vanguard Press, Buckingham, UK
137. Seddon, J., O'donovan, B., and Zokaei, K., *Rethinking Lean Service*, in *Service Design and Delivery*. 2011, Springer. p. 41-60.
138. Seddon, J., *Lean Is a Waning Fad.* Management Services, 2011. 55(4), p. 34-36.
139. Radnor, Z., *Transferring Lean into Government.* Journal of Manufacturing Technology Management, 2010.
140. Lawal, A.K., Rotter, T., Kinsman, L., Sari, N., Harrison, L., Jeffery, C., et al., *Lean Management in Health Care: Definition, Concepts,*

Methodology and Effects Reported (Systematic Review Protocol). Systematic reviews, 2014. 3(1), p. 1-6.
141. Radnor, Z.J., Holweg, M., and Waring, J., *Lean in Healthcare: The Unfilled Promise?* Social Science & Medicine, 2012. 74(3), p. 364-371.
142. Antony, J., Sreedharan, R., Chakraborty, A., and Gunasekaran, A., *A Systematic Review of Lean in Healthcare: A Global Prospective.* International Journal of Quality & Reliability Management, 2019.
143. Preston, N.J., *The Health of the Nation Outcome Scales: Validating Factorial Structure and Invariance across Two Health Services.* Australian & New Zealand Journal of Psychiatry, 2000. 34(3), p. 512-519.
144. Mason, S., Nicolay, C., and Darzi, A., *The Use of Lean and Six Sigma Methodologies in Surgery: A Systematic Review.* The Surgeon, 2015. 13(2), p. 91-100.
145. Burgess, N. and Radnor, Z., *Evaluating Lean in Healthcare.* International Journal of Health Care Quality Assurance, 2013. 26(3), p. 220 - 235.
146. Weimarsson, H., *Nine out of Ten Hospitals Changed to" Lean".* Lakartidningen, 2011. 108(39), p. 1915-1915.
147. Institutet, K. *Website of the Karolinska Institutet*. 2020 [cited 2020 2020]; Available from: https://ki.se/en.
148. Ker, J.-I., Wang, Y., Hajli, M.N., Song, J., and Ker, C.W., *Deploying Lean in Healthcare: Evaluating Information Technology Effectiveness in Us Hospital Pharmacies.* International Journal of Information Management, 2014. 34(4), p. 556-560.
149. Hartmann, D.H. *Pharmacy Optimisation: Getting People on Drugs Faster*. in *International Conference on Industrial Engineering, Systems Engineering and Engineering Management*. 2011. Stellenbosch.
150. Kachwee, M. and Hartmann, D. *Hospital Supply Chain Management and Optimisation*. in *Southern African Institute for Industrial Engineering (SAIIE) 25*. 2013. Stellenbosch.
151. Ryan, A., Hunter, K., Cunningham, K., Williams, J., O'shea, H., Rooney, P., et al., *Steps: Lean Thinking, Theory of Constraints and Identifying Bottlenecks in an Emergency Department.* Irish Medical Journal, 2013. 106(4), p. 105-7.

152. Hartmann, D. and Mandavha, R. *Lean Healthcare, a Casualty of Inefficiency*. in *SAIIE Conference*. 2010. Pretoria, Southern African Institute for Industrial Engineering.
153. Nicolay, C., Purkayastha, S., Greenhalgh, A., Benn, J., Chaturvedi, S., Phillips, N., et al., *Systematic Review of the Application of Quality Improvement Methodologies from the Manufacturing Industry to Surgical Healthcare*, in *Database of Abstracts of Reviews of Effects (Dare): Quality-Assessed Reviews [Internet]*. 2012, Centre for Reviews and Dissemination (UK).
154. Hartmann, D. and Sunjka, B., *Private Theatre Utilisation in South Africa: A Case Study.* South African Medical Journal, 2013. 103(5), p. 285-7.
155. Hartmann, D. and Sooklal, S., *The Pen Is Mightier Than the Scalpel : The Case for Electronic Medical Records : Case Study.* South African Journal of Industrial Engineering, 2012. 23(2), p. 191 - 201.
156. Silvester, K., Lendon, R., Bevan, H., Steyn, R., and Walley, P., *Reducing Waiting Times in the Nhs: Is Lack of Capacity the Problem?* Clinician in Management, 2004. 12(3), p. 105-109.
157. Kenney, C., *Transforming Health Care: Virginia Mason Medical Center's Pursuit of the Perfect Patient Experience*. 2012, CRC Press, Boca Raton, Florida
158. Weber, D.O., *Toyota-Style Management Drives Virginia Mason.* Physician Executive, 2006. 32(1), p. 12.
159. Meredith, J.O., Grove, A.L., Walley, P., Young, F., and Macintyre, M.B., *Are We Operating Effectively? A Lean Analysis of Operating Theatre Changeovers.* Operations Management Research, 2011. 4(3-4), p. 89.
160. Scherrer-Rathje, M., Boyle, T.A., and Deflorin, P., *Lean, Take Two! Reflections from the Second Attempt at Lean Implementation.* Business horizons, 2009. 52(1), p. 79-88.
161. Mazzocato, P., Savage, C., Brommels, M., Aronsson, H., and Thor, J., *Lean Thinking in Healthcare: A Realist Review of the Literature.* BMJ Quality & Safety, 2010. 19(5), p. 376-382.
162. Lucey, J., Bateman, N., and Hines, P., *Why Major Lean Transitions Have Not Been Sustained.* Management Services, 2005. 49(2), p. 9-13.
163. Andersen, H., Røvik, K.A., and Ingebrigtsen, T., *Lean Thinking in Hospitals: Is There a Cure for the Absence of Evidence? A Systematic Review of Reviews.* BMJ open, 2014. 4(1).

164. Glasgow, J.M., Scott-Caziewell, J.R., and Kaboli, P.J., *Guiding Inpatient Quality Improvement: A Systematic Review of Lean and Six Sigma.* The Joint Commission Journal on Quality and Patient Safety, 2010. 36(12), p. 533-AP5.
165. Brandão De Souza, L., *Trends and Approaches in Lean Healthcare.* Leadership in Health Services, 2009. 22(2), p. 121-139.
166. Alexander, J.A. and Hearld, L.R., *What Can We Learn from Quality Improvement Research? A Critical Review of Research Methods.* Medical Care Research and Review, 2009. 66(3), p. 235-271.
167. Radnor, Z. and Boaden, R., *Editorial: Lean in Public Services—Panacea or Paradox?* Public Money & Management, 2008. 28(1), p. 3 - 7.
168. Leseure, M., Hudson-Smith, M., and Radnor, Z., *Transferring Lean into Government.* Journal of Manufacturing Technology Management, 2010.
169. Kim, C.S., Spahlinger, D.A., Kin, J.M., and Billi, J.E., *Lean Health Care: What Can Hospitals Learn from a World-Class Automaker?* Journal of Hospital Medicine, 2006. 1(3), p. 191-199.
170. Jones, D. and Mitchell, A., *Lean Thinking for the Nhs.* London: NHS Confederation, 2006.
171. Rotter, T., Plishka, C., Lawal, A., Harrison, L., Sari, N., Goodridge, D., et al., *What Is Lean Management in Health Care? Development of an Operational Definition for a Cochrane Systematic Review.* Evaluation & the Health Professions, 2019. 42(3), p. 366-390.
172. Sloan, T., Fitzgerald, A., Hayes, K.J., Radnor, Z., Robinson, S., Sohal, A., et al., *Complexity Complicates Lean: Lessons from Seven Emergency Services.* Journal of Health Organization and Management, 2014. 28(2), p. 266 - 288.
173. Fujimoto, T., *The Evolution of a Manufacturing System at Toyota*. 1999, Oxford University Press, Oxford, UK
174. Hines, P., Holweg, M., and Rich, N., *Learning to Evolve: A Review of Contemporary Lean Thinking.* International Journal of Operations & Production Management, 2004. 24(10), p. 994-1011.
175. Mazzocato, P., Holden, R.J., Brommels, M., Aronsson, H., Bäckman, U., Elg, M., et al., *How Does Lean Work in Emergency Care? A Case Study of a Lean-Inspired Intervention at the Astrid Lindgren Children's Hospital, Stockholm, Sweden.* BMC Health Services Research, 2012. 12(1), p. 1-13.

176. Plsek, P.E. and Greenhalgh, T., *Complexity Science: The Challenge of Complexity in Health Care.* BMJ: British Medical Journal, 2001. 323(7313), p. 625.
177. Plsek, P.E. and Wilson, T., *Complexity, Leadership, and Management in Healthcare Organisations.* BMJ: British Medical Journal, 2001. 323(7315), p. 746.
178. Burgess, N. and Radnor, Z., *Thesis: Evaluating Lean in Healthcare*. 2012, University of Warwick, Warwick, UK.
179. Ingebrigtsen, T., *Merging of Three Hospitals into One University Hospital 813–7.*
180. Proudlove, N., Moxham, C., and Boaden, R., *Lessons for Lean in Healthcare from Using Six Sigma in the Nhs.* Public Money & Management, 2008. 28(1), p. 27-34.
181. Trist, E., *The Evolution of Socio-Technical Systems.* Occasional Paper, 1981. 2, p. 1981.
182. Walshe, K., *Understanding What Works—and Why—in Quality Improvement: The Need for Theory-Driven Evaluation.* International Journal for Quality in Health Care, 2007. 19(2), p. 57-59.
183. Mcnulty, T. and Ferlie, E., *Reengineering Health Care: The Complexities of Organizational Transformation*. 2002, OUP Oxford, Oxford, UK
184. Seddon, J., *Discussion of Failure Demand Repeat Consumption and the Idea of Demand*, D. Hartmann, Editor. 2020, Johannesburg, South Africa.
185. Marr, B. and Neely, A., *Managing and Measuring for Value: The Case of Call Centre Performance*. 2004, Cranfield School of Management and Fujitsu, Cranfield, UK.
186. Walley, P., *Does the Public Sector Need a More Demand-Driven Approach to Capacity Management?* Production Planning & Control, 2013. 24(10-11), p. 877-890.
187. O'sullivan, A. and Sheffrin, S.M., *Economics: Principles in Action*. 2007, Pearson/Prentice Hall, Upper Saddle River, New Jersey
188. Mankiw, N.G., *Principles of Economics*. 2014, Nelson Education, Toronto, Canada
189. Demir, E., Gunal, M.M., and Southern, D., *Demand and Capacity Modelling for Acute Services Using Discrete Event Simulation.* Health Systems, 2017. 6(1), p. 33-40.
190. Allder, S., Silvester, K., and Walley, P., *Managing Capacity and Demand across the Patient Journey.* 2010. 10(1), p. 13.

191. Lovelock, C.H., *Strategies for Managing Demand in Capacity-Constrained Service Organisations.* The Service Industries Journal, 1984. 4(3), p. 12-30.
192. Allder, S., Walley, P., and Silvester, K., *Is Follow-up Capacity the Current Nhs Bottleneck?* Clinical Medicine, 2011. 11(1), p. 31.
193. Walley, P. and Adams, M.M., *An Evaluation of Demand Management Practices in Uk Police Forces.* 2019.
194. Walley, P., Found, P., and Williams, S., *Failure Demand: A Concept Evaluation in Uk Primary Care.* International Journal of Health Care Quality Assurance, 2019. 32(1), p. 21-33.
195. Praeter, J., *Queues in Health.* Health Care Management Science, 2002. 5(4), p. 283 - 283.
196. Hartmann, D. *I Just Don't Have Enough Time*. in *Southern African Institute for Industrial Engineering*. 2010. Johannesburg.
197. Toussaint, J. and Gerard, R., *On the Mend: Revolutionizing Healthcare to Save Lives and Transform the Industry*. 2010, Lean Enterprise Institute, Cambridge, Massachusetts
198. Wagstaff, A., Van Doorslaer, E., and Paci, P., *On the Measurement of Horizontal Inequity in the Delivery of Health Care.* Journal of Health Economics, 1991. 10, p. 169-205.
199. Crandall, R.E. and Markland, R.E., *Demand Management-Today's Challenge for Service Industries.* Production and Operations Management, 1996. 5(2), p. 106-120.
200. Hosseini, H., *Understanding the Market Mechanism before Adam Smith: Economic Thought in Medieval Islam.* History of Political Economy, 1995. 27(3), p. 539-561.
201. Marshall, A., *Principles of Economics*. 2004, Digireads. com Publishing
202. Macario, A., *Are Your Operating Theatres Being Run Efficiently?* Medscape Anesthesiology, 2010.
203. Schoen, C., Osborn, R., Doty, M.M., Bishop, M., Peugh, J., and Murukutla, N., *Toward Higher-Performance Health Systems: Adults' Health Care Experiences in Seven Countries, 2007.* Health Aff (Millwood), 2007. 26(6), p. w717-34.
204. South African National Department of Health, *National Department of Health Strategic Plan 2010/2011 - 2012/2013*, S.A.N.D.o. Health, Editor. 2010, Pretoria. p. 115.

205. South African National Department of Health, *Strategic Plan 2014/15 - 2018/19*. 2014, South African National Department of Health,, Pretoria, South Africa.
206. Louis M. Savary, C.C.-M., *The Nun and the Bureaucrat: How They Found an Unlikely Cure for America's Sick Hospitals*. 2006, CC-M Productions, Washington
207. Hartmann, D., *Improvement in Operating Theatre Efficiency through Better Measurement and Scheduling Msc. Research Report*. 2012, University of the Witwatersrand Press, Johannesburg.
208. Heskett, J.L., *Service Breakthroughs*. 1990, Simon and Schuster, New York City
209. Klassen, K.J. and Rohleder, T.R., *Combining Operations and Marketing to Manage Capacity and Demand in Services.* Service Industries Journal, 2001. 21(2), p. 1-30.
210. Jack, E.P. and Powers, T.L., *A Review and Synthesis of Demand Management, Capacity Management and Performance in Health-Care Services.* International Journal of Management Reviews, 2009. 11(2), p. 149-174.
211. Hopp, W.J. and Lovejoy, W.S., *Hospital Operations: Principles of High Efficiency Health Care*. 2012, FT Press, Upper Saddle River, New Jersey
212. Naidoo, J. and Wills, J., *Health Promotion: Foundations for Practice*. 2010, Baillière Tindall, Sidcup, UK
213. Hogg, W., Rowan, M., Russell, G., Geneau, R., and Muldoon, L., *Framework for Primary Care Organizations: The Importance of a Structural Domain.* International Journal for Quality in Health Care, 2008. 20(5), p. 308-313.
214. World Health Organization, *Bangladesh Health System Review*. 2015, Manila: WHO Regional Office for the Western Pacific
215. Maeda, A., Araujo, E., Cashin, C., Harris, J., Ikegami, N., and Reich, M.R., *Universal Health Coverage for Inclusive and Sustainable Development: A Synthesis of 11 Country Case Studies*. 2014, The World Bank, Washington D.C.
216. El-Saharty, S., Sparkes, S.P., Barroy, H., Ahsan, K.Z., and Ahmed, S.M., *The Path to Universal Health Coverage in Bangladesh: Bridging the Gap of Human Resources for Health*. 2015, The World Bank, Washington D.C.

217. Unctad. *Un List of Least Developed Countries*. 2020 [cited 2020; Available from: https://unctad.org/topic/vulnerable-economies/least-developed-countries/list.
218. Osborn, R., Squires, D., Doty, M.M., Sarnak, D.O., and Schneider, E.C., *In New Survey of Eleven Countries, Us Adults Still Struggle with Access to and Affordability of Health Care.* Health Affairs, 2016. 35(12), p. 2327-2336.
219. Richardson, J., *Supply and Demand for Medical Care: Or, Is the Health Care Market Perverse?* Australian Economic Review, 2001. 34(3), p. 336-352.
220. Tajeu, G., *Health Maintenance Organization (Hmo)*, in *The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society*. 2014, John Wiley & Sons, Ltd, Hoboken, New Jersey.
221. Heineke, J., *Strategic Operations Management Decisions and Professional Performance in Us Hmos.* Journal of Operations Management, 1995. 13(4), p. 255-272.
222. Jack, E.P. and Powers, T.L., *Volume Flexible Strategies in Health Services: A Research Framework.* Production and Operations Management, 2004. 13(3), p. 230-244.
223. Lewin, M.E. and Altman, S., *America's Health Care Safety Net: Intact but Endangered*. 2000, National Academies Press, Washington D.C.
224. World Health Organization. *Naming the Coronavirus Disease (Covid-19) and the Virus That Causes It*. 2020; Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it.
225. The Straits Times, *Coronavirus: Wuhan to Build 19 More Makeshift Hospitals as China Struggles to Contain Epidemic*, in *The straits times*. 2020, Singapore Press Holdings, Singapore, Singapore.
226. World Health Organization, *Who Director-General's Opening Remarks at the Media Briefing on Covid-19 - 11 March 2020*. 2020, World Health Organization, .
227. Schuler, R., *60 Bis 70 Prozent in Deutschland Werden Sich Infizieren*, in *Bild*. 2020, Axel Springer SE, Berlin, Germany.
228. Regalado, A., *Worst-Case Coronavirus Scenario: 214 Million Americans Infected, 1.7 Million Dead*, in *MIT Technology Review*. 2020, MIT Technology Review Cambridge Massachusetts.
229. Qualls, N., Levitt, A., Kanade, N., Wright-Jegede, N., Dopson, S., Biggerstaff, M., et al., *Community Mitigation Guidelines to Prevent*

Pandemic Influenza—United States, 2017. MMWR Recommendations and Reports, 2017. 66(1), p. 1.
230. World Health Organization. *Coronavirus Disease (Covid-19) Advice for the Public*. 2020 [cited 2020; Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public.
231. Lorig, K.R., Mazonson, P.D., and Holman, H.R., *Evidence Suggesting That Health Education for Self-Management in Patients with Chronic Arthritis Has Sustained Health Benefits While Reducing Health Care Costs.* Arthritis & Rheumatology, 1993. 36(4), p. 439-446.
232. Donnelley, R., *Mental Health Collaborative Programme Toolkit*, T.S. Government, Editor. 2009, Crown, Edinburgh.
233. Healthcare Improvement Scotland, *A Guide to Understanding and Managing Demand*. 2019, Mental Health Access Improvement Support Team (MHAIST), Edinburgh, UK.
234. El-Ghorr, A., Cameron, R., Fleming, M., Mckechnie, L., Thomson, D., and Doherty, S., *Scotland's National Approach to Improving Mental Health Services: Integrated Care Pathways as Tools for Redesign and Continuous Quality Improvement.* 2010. 14(2), p. 57-64.
235. Seddon, J. and Brand, C., *Debate: Systems Thinking and Public Sector Performance.* 2008.
236. Donaldson, C. and Gerard, K., *Paying General Practitioners: Shedding Light on the Review of Health Services.* Journal of the Royal College of General Practitioners, 1989. 39(320), p. 114-117.
237. Richardson, J.R. and Peacock, S.J., *Supplier-Induced Demand.* Applied Health Economics and Health Policy, 2006. 5(2), p. 87-98.
238. United States Department of Health Human Services, *Healthy People 2000: National Health Promotion and Disease Prevention Objectives*, in *Healthy People 2000: National Health Promotion and Disease Prevention Objectives*. 1991, US Government Printing Office.
239. Deming, E., *Out of the Crisis. Quality, Productivity and Competitive Position. Cambridge, Ma: Center for Advanced Engineering Study*. 1986, Cambridge University Press.
240. Masters, K.I.A., *From New Public Management to Lean Thinking: Understanding and Managing'potentially Avoidable Failure Induced Demand'*. 2010, University of Sussex.

241. Lillrank, P., Groop, P.J., and Malmström, T.J., *Demand and Supply–Based Operating Modes—a Framework for Analyzing Health Care Service Production.* Milbank Quarterly, 2010. 88(4), p. 595-615.
242. Hm Government, *National Indicators for Local Authorities and Local Authority Partnerships: Handbook of Definitions*, C. Office, Editor. 2007, DCLG, London, UK.
243. Hm Government, *Ni 014 - Avoidable Contact:*. 2010, Cabinet Office, London, UK.
244. Seddon, J., *Ni 14 ('Avoidable Contact'):–from the Horse's Mouth'.* Customer Strategy, 2009. 2(1), p. 33-34.
245. Martin, P., *'Avoidable Contact: If Only....* 2017, Local Government Chronicle.
246. Clay, H. and Stern, R., *Making Time in General Practice.* Birmingham: NHS Alliance, 2015, p. 59-63.
247. Mccartney, M., *Pointless Paperwork, Not Patients, Is What Gps Should Avoid.* BMJ : British Medical Journal (Online), 2017. 359.
248. Armstrong, D.M., *Belief, Truth and Knowledge*. 1973, CUP Archive
249. Seth, A., *Epistemology and Ontology.* The Philosophical Review, 1894. 3(5), p. 568-582.
250. Armstrong, D.M. *A Naturalist Program: Epistemology and Ontology*. in *Proceedings and Addresses of the American Philosophical Association*. 1999. Newark, Delaware, JSTOR.
251. Raddon, A., *Early Stage Research Training: Epistemology & Ontology in Social Science Research.* Generic Skills Training for Research Students, 2010, p. 1-14.
252. May, T. and Williams, M., *Knowing the Social World*. 1998, Open University Press, Maidenhead, UK
253. Berryman, D.R., *Ontology, Epistemology, Methodology, and Methods: Information for Librarian Researchers.* Medical Reference Services Quarterly, 2019. 38(3), p. 271-279.
254. Al-Saadi, H., *Demystifying Ontology and Epistemology in Research Methods.* Research Gate, 2014. 1(1), p. 1-10.
255. Saunders, M., Lewis, P., and Thornhill, A., *Research Methods for Business Students*. 8th ed. 2019, Pearson Education Ltd, Harlow, United Kingdom
256. Moon, K. and Blackman, D., *A Guide to Ontology, Epistemology, and Philosophical Perspectives for Interdisciplinary Researchers*. 2017.

257. Goertz, G. and Mahoney, J., *Concepts and Measurement: Ontology and Epistemology.* Social Science Information, 2012. 51(2), p. 205-216.
258. Cresswell, J.W., *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. 2009, Sage, Thousand Oaks
259. Lawson, T., *Comparing Conceptions of Social Ontology: Emergent Social Entities and/or Institutional Facts?* Journal for the Theory of Social Behaviour, 2016. 46(4), p. 359-399.
260. Moon, K. and Blackman, D., *A Guide to Understanding Social Science Research for Natural Scientists.* Conservation Biology, 2014. 28(5), p. 1167-1177.
261. Ahmed, A., *Ontological, Epistemological and Methodological Assumptions: Qualitative Versus Quantitative.* Research Gate, 2008.
262. Moodley, K. and Myer, L., *Health Research Ethics Committees in South Africa 12 Years into Democracy.* BMC Medical Ethics, 2007. 8(1), p. 1.
263. Norton, J.D., *Are Thought Experiments Just What You Thought?* Canadian Journal of Philosophy, 1996. 26(3), p. 333-366.
264. De Grasse Tyson, N., *Scientific Thinking*, in *Neil De Grasse Tyson Teaches Scientific Thinking and Communication Skills*. 2019, Yanka Industries,, San Francisco USA.
265. Sorensen, R.A., *Thought Experiments*. 1998, Oxford University Press on Demand, Oxford, UK
266. Norton, J., *Thought Experiments in Einstein's Work.* Horowitz and Massey, 1991. 1991, p. 129-148.
267. Brown, J.R., *Review of Thought Experiments by Roy Sorensen.* Canadian Journal of Philosophy, 1995. 25(1), p. 135-142.
268. Brown, J.R., *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*. 2011, Routledge, New York City
269. Di Paolo, E.A., Noble, J., and Bullock, S., *Simulation Models as Opaque Thought Experiments.* Univerisity of Southampton, 2000.
270. Forrester, J.W., *Counterintuitive Behavior of Social Systems.* Theory and decision, 1971. 2(2), p. 109-140.
271. Jabareen, Y., *Building a Conceptual Framework: Philosophy, Definitions, and Procedure.* International Journal of qualitative Methods, 2009. 8(4), p. 49-62.
272. Sterman, J.D., *Business Dynamics: Systems Thinking and Modeling for a Complex World*. 2000, McGraw Hill, Homewood, Illinois

273. Sterman, J., *Business Dynamics*. 2010, Irwin/McGraw-Hill, Homewood, Illinois
274. Brzezina, N., Kopainsky, B., and Mathijs, E., *Can Organic Farming Reduce Vulnerabilities and Enhance the Resilience of the European Food System? A Critical Assessment Using System Dynamics Structural Thinking Tools.* Sustainability, 2016. 8(10), p. 971.
275. Haraldsson, H.V., *Introduction to Systems and Causal Loop Diagrams.* System Dynamic Course, Lumes, Lund University, Sweden, 2000.
276. Bulsara, C., *Using a Mixed Methods Approach to Enhance and Validate Your Research.* Brightwater Group Research Centre, 2015, p. 1-82.
277. Womack, J., *Gemba Walks: Expanded 2nd Edition*. 2013, Lean Enterprise Institute, Cambridge, Massachusetts
278. Roser, C., Lorentzen, K., and Deuse, J., *Reliable Shop Floor Bottleneck Detection for Flow Lines through Process and Inventory Observations: The Bottleneck Walk.* Logistics Research, 2015. 8(1), p. 7.
279. Goodson, R.E., *Read a Plant-Fast.* Harvard Business Review, 2002. 80(5), p. 105-113.
280. Hopp, W.J., *Single Server Queueing Models*, in *Building Intuition*. 2008, Springer. p. 51-79.
281. Jorgensen, D.L., *Participant Observation.* Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource, 2015, p. 1-15.
282. Jersild, A.T. and Meigs, M.F., *Chapter V: Direct Observation as a Research Method.* Review of Educational Research, 1939. 9(5), p. 472-482.
283. Iacono, J., Brown, A., and Holtham, C., *Research Methods--a Case Example of Participant Observation.* Electronic Journal of Business Research Methods, 2009. 7(1).
284. Madden, S., *From Databases to Big Data.* IEEE Internet Computing, 2012. 16(3), p. 4-6.
285. De Mauro, A., Greco, M., and Grimaldi, M. *What Is Big Data? A Consensual Definition and a Review of Key Research Topics*. in *AIP conference proceedings*. 2015. American Institute of Physics.
286. Sengupta, C., *Financial Modeling Using Excel and Vba*. Vol. 152. 2004, John Wiley & Sons, Hoboken, New Jersey

287. Guttag, J., *Introduction to Computation and Programming Using Python: With Application to Understanding Data*. 2016, MIT Press
288. Bird, J., *Science for Engineering*. 2015, Routledge
289. Zelle, J.M., *Python Programming: An Introduction to Computer Science*. 2004, Franklin, Beedle & Associates, Inc.
290. Shaw, Z.A., *Learn Python 3 the Hard Way: A Very Simple Introduction to the Terrifyingly Beautiful World of Computers and Code*. 2017, Addison-Wesley Professional
291. Barth, A., Caillaud, E., and Rose, B. *How to Validate Research in Engineering Design?* in *DS 68-2: Proceedings of the 18th International Conference on Engineering Design (ICED 11), Impacting Society through Engineering Design, Vol. 2: Design Theory and Research Methodology, Lyngby/Copenhagen, Denmark, 15.-19.08. 2011*. 2011.
292. Rosenthal, R. and Rosnow, R.L., *Essentials of Behavioral Research: Methods and Data Analysis*. Vol. 2. 1991, McGraw-Hill New York
293. Bracht, G.H. and Glass, G.V., *The External Validity of Experiments.* American Educational Research Journal, 1968. 5(4), p. 437-474.
294. Shanmuganathan, T., *Ethics and the Observer's Paradox.* relc Journal, 2005. 36(1), p. 73-84.
295. Wickström, G. and Bendix, T., *The" Hawthorne Effect"—What Did the Original Hawthorne Studies Actually Show?* Scandinavian Journal of Work, Environment & Health, 2000, p. 363-367.
296. Uysal, M., *The Determinants of Tourism Demand*. The Economic Geography of the Tourist Industry: A Supply-Side Analysis. Vol. 79. 1998
297. Balogun, O. and Yusuf, S., *Determinants of Demand for Microcredit among the Rural Households in South-Western States, Nigeria.* Journal of Agriculture and Social Sciences - Pakistan, 2011.
298. Kim, J.-H. and Park, D., *The Determinants of Demand for Private Tutoring in South Korea.* Asia Pacific Education Review, 2010. 11(3), p. 411-421.
299. Vecino-Ortiz, A.I., *Determinants of Demand for Antenatal Care in Colombia.* Health Policy, 2008. 86(2-3), p. 363-372.
300. Ravitch, S.M. and Riggan, M., *Reason & Rigor: How Conceptual Frameworks Guide Research*. 2016, Sage Publications, Newbury Park, California

301. Collier, D. and Levitsky, S., *Democracy with Adjectives: Conceptual Innovation in Comparative Research.* World Politics, 2011. 49(3), p. 430-451.
302. Häggqvist, S., *A Model for Thought Experiments.* Canadian Journal of Philosophy, 2009. 39(1), p. 55-76.
303. Gendler, T.S., *Thought Experiment: On the Powers and Limits of Imaginary Cases*. 2014, Routledge, Abingdon, UK
304. Stoves, J., Connolly, J., Cheung, C.K., Grange, A., Rhodes, P., O'donoghue, D., et al., *Electronic Consultation as an Alternative to Hospital Referral for Patients with Chronic Kidney Disease: A Novel Application for Networked Electronic Health Records to Improve the Accessibility and Efficiency of Healthcare.* Qual Saf Health Care, 2010. 19(5), p. e54-e54.
305. World Health Organization, *Constitution*. 1989, World Health Organization, Geneva, Switzerland
306. Havemann, R. and Van Der Berg, S., *The Demand for Health Care in South Africa.* Journal for the Study of Economics and Econometrics, 2003. 27(3), p. 1-27.
307. Moerman, D.E., *Meaning, Medicine, and the" Placebo Effect"*. Vol. 28. 2002, Cambridge University Press Cambridge
308. Borossa, J. and Rooney, C., *Fortress Hypochondria: Health and Safety*, in *Nationalism and the Body Politic*. 2018, Routledge. p. 5-19.
309. Rogers, A., Hassell, K., and Nicolaas, G., *Demanding Patients?: Analysing the Use of Primary Care*. 1999, Open University Press, Maidenhead, UK
310. Grossman, M., *On the Concept of Health Capital and the Demand for Health.* Journal of Political Economy, 1972. 80(2), p. 223-255.
311. Stevens, B.H. and Rydell, C.P., *Spatial Demand Theory and Monopoly Price Policy.* Papers in Regional Science, 1966. 17(1), p. 195-204.
312. World Health Organization, *World Health Statistics 2014.* 2014.
313. Almeida, A.S. and Cima, J.F., *Demand Uncertainty and Hospital Costs: An Application to Portuguese Public Hospitals.* The European Journal of Health Economics, 2015. 16(1), p. 35-45.
314. Brailsford, S.C., Lattimer, V., Tarnaras, P., and Turnbull, J., *Emergency and on-Demand Health Care: Modelling a Large Complex System.* Journal of the Operational Research Society, 2004, p. 34-42.

315. Chapple, A. and Rogers, A., *'Self-Care' and Its Relevance to Developing Demand Management Strategies: A Review of Qualitative Research.* Health & Social Care in the Community, 1999. 7(6), p. 445-454.
316. Gallagher, M., Pearson, P., Drinkwater, C., and Guy, J., *Managing Patient Demand: A Qualitative Study of Appointment Making in General Practice.* The British Journal of General Practice, 2001. 51(465), p. 280-285.
317. Gillam, S. and Pencheon, D., *Managing Demand in General Practice.* BMJ : British Medical Journal, 1998. 316(7148), p. 1895-1898.
318. Fries, J.F., Koop, C.E., Beadle, C.E., Cooper, P.P., England, M.J., Greaves, R.F., et al., *Reducing Health Care Costs by Reducing the Need and Demand for Medical Services.* New England Journal of Medicine, 1993. 329(5), p. 321-325.
319. Goldberg, S.E., *Demand Management: Implementing Your Own Program.* Fam Pract Manag, 1998. 5(8), p. 49-50, 55-6, 59-62.
320. Seddon, J. and Caulkin, S., *Systems Thinking, Lean Production and Action Learning.* Action Learning: Research and Practice, 2007. 4(1), p. 9-24.
321. Wagstaff, A., *The Demand for Health: Some New Empirical Evidence.* Journal of Health Economics, 1986. 5(3), p. 195-233.
322. Brennan, T.J., *Demand-Side Management Programs under Retail Electricity Competition*. 1998, Resources for the Future Washington, DC
323. Yu, D., Lowndes, B., Thiels, C., Bingener, J., Abdelrahman, A., Lyons, R., et al., *Quantifying Intraoperative Workloads across the Surgical Team Roles: Room for Better Balance?* World Journal of Surgery, 2016. 40(7), p. 1565-1574.
324. Connelly, J.E., Philbrick, J.T., Smith Jr, G.R., Kaiser, D.L., and Wymer, A., *Health Perceptions of Primary Care Patients and the Influence on Health Care Utilization.* Medical care, 1989, p. S99-S109.
325. Wennberg, J.E., *Outcomes Research, Cost Containment, and the Fear of Health Care Rationing*. 1990, Massachusetts Medical Society, Waltham, Massachusetts.
326. Domenighetti, G., Casabianca, A., Gutzwiller, F., and Martinoli, S., *Revisiting the Most Informed Consumer of Surgical Services: The Physician-Patient.* International Journal of Technology Assessment in Health Care, 1993. 9(4), p. 505-513.

327. Donaldson, C. and Gerard, K., *Countering Moral Hazard in Public and Private Health Care Systems: A Review of Recent Evidence.* Journal of Social Policy, 1989. 18(2), p. 235-251.
328. Horne, R., James, D., Petrie, K., Weinman, J., and Vincent, R., *Patients' Interpretation of Symptoms as a Cause of Delay in Reaching Hospital During Acute Myocardial Infarction.* Heart, 2000. 83(4), p. 388-393.
329. Leslie, W., Urie, A., Hooper, J., and Morrison, C., *Delay in Calling for Help During Myocardial Infarction: Reasons for the Delay and Subsequent Pattern of Accessing Care.* Heart, 2000. 84(2), p. 137-141.
330. Storla, D.G., Yimer, S., and Bjune, G.A., *A Systematic Review of Delay in the Diagnosis and Treatment of Tuberculosis.* BMC Public Health, 2008. 8(1), p. 15.
331. Van Der Werf, M., Chechulin, Y., Yegorova, O., Marcinuk, T., Stopolyanskiy, A., Voloschuk, V., et al., *Health Care Seeking Behaviour for Tuberculosis Symptoms in Kiev City, Ukraine.* The International Journal of Tuberculosis and Lung Disease, 2006. 10(4), p. 390-395.
332. Molloy, G.J., Mcgee, H.M., O'neill, D., and Conroy, R.M., *Loneliness and Emergency and Planned Hospitalizations in a Community Sample of Older Adults.* Journal of the American Geriatrics Society, 2010. 58(8), p. 1538-1541.
333. Brailsford, S. and Hilton, N., *A Comparison of Discrete Event Simulation and System Dynamics for Modelling Health Care Systems*. 2001.
334. Roberts, C. and Dangerfield, B., *Modelling the Epidemiological Consequences of Hiv Infection and Aids: A Contribution from Operational Research.* Journal of the Operational Research Society, 1990. 41(4), p. 273-289.
335. Rauner, M. *Managing the Aids Epidemic in Vienna, Austria: Prevention Strategies for the 21st Century*. in *Information, Management and Planning of Health Services, Proceedings of the 25th Meeting of the European Working Group on Operational Research Applied to Health Services*. 1999.
336. Davies, R. and Davies, H., *Modelling Patient Flows and Resource Provision in Health Systems.* Omega, 1994. 22(2), p. 123-131.
337. Bowen, S. and Zwi, A.B., *Pathways to "Evidence-Informed" Policy and Practice: A Framework for Action.* PLoS Med, 2005. 2(7).

338. Ahmed, S., Begum, T., and Cotlear, D., *Bangladesh: Unravelling the 'Good Health at Low Cost' Story*. 2019, World Bank Group, Washington D.C.
339. Levine, R.J., *Institutional Review Boards.* BMJ: British Medical Journal, 1989. 298(6683), p. 1268.
340. World Medical Association, *World Medical Association Declaration of Helsinki. Ethical Principles for Medical Research Involving Human Subjects.* Bulletin of the World Health Organization, 2001. 79(4), p. 373.
341. Wallenius, J.J., *Failure Demand at Telefonbanken, Seb*. 2011, Chalmers University of Technology, Göteborg, Sweden.
342. Piercy, N. and Rich, N., *High Quality and Low Cost: The Lean Service Centre.* European Journal of Marketing, 2009. 43(11/12), p. 1477-1497.
343. Pidgeon, C., *Systems Thinking and Public Sector Efficiencies*. 2010, Northern Ireland Assembly, Research and Library Service, Northern Ireland.
344. Plsek, P.E. and Greenhalgh, T., *The Challenge of Complexity in Health Care.* BMJ: British Medical Journal, 2001. 323(7313), p. 625.
345. Kohn, L.T., Corrigan, J.M., and Donaldson, M.S., *To Err Is Human: Building a Safer Health System*. Vol. 627. 2000, National Academies Press, Washington D.C.
346. Sox, H.C., *Medical Decision Making*. 1988, ACP Press, Sydney, Australia
347. Hopp, W.J., *Supply Chain Science*. 2011, Waveland Press, Long Grove, Illinois
348. Mccarthy, D. and Rich, N., *Lean Tpm: A Blueprint for Change*. 2015, Butterworth-Heinemann, Oxford, UK
349. Graeber, D., *On the Phenomenon of Bullshit Jobs: A Work Rant.* Strike Magazine, 2013. 15(3), p. 1-5.
350. Dijkstra, J., Galbraith, R., Hodges, B.D., Mcavoy, P.A., Mccrorie, P., Southgate, L.J., et al., *Expert Validation of Fit-for-Purpose Guidelines for Designing Programmes of Assessment.* BMC Medical Education, 2012. 12(1), p. 1-8.
351. Hans, E.W., Van Houdenhoven, M., and Hulshof, P.J., *A Framework for Healthcare Planning and Control*, in *Handbook of Healthcare System Scheduling*. 2012, Springer. p. 303-320.
352. Frenk, J., *Dimensions of Health System Reform.* Health Policy, 1994. 27(1), p. 19-34.

353. World Health Organization, *The World Health Report 2000: Health Systems: Improving Performance*. 2000, World Health Organization
354. Sempere-Selva, T., Peiró, S., Sendra-Pina, P., Martínez-Espín, C., and López-Aguilera, I., *Inappropriate Use of an Accident and Emergency Department: Magnitude, Associated Factors, and Reasons—an Approach with Explicit Criteria.* Annals of Emergency Medicine, 2001. 37(6), p. 568-579.
355. World Health Organization. *Noncommunicable Diseases*. Health Topics 2015 2017-06-26 [cited 2017 2017-06-26].
356. Bernell, S. and Howard, S.W., *Use Your Words Carefully: What Is a Chronic Disease?* Frontiers in Public Health, 2016. 4, p. 159.
357. Wagner, E.H., Grothaus, L.C., Sandhu, N., Galvin, M.S., Mcgregor, M., Artz, K., et al., *Chronic Care Clinics for Diabetes in Primary Care.* A system-wide Randomized Trial, 2001. 24(4), p. 695-700.
358. Wennberg, J.E., *Dealing with Medical Practice Variations: A Proposal for Action.* Health Affairs, 1984. 3(2), p. 6-32.
359. Das, P. and Horton, R., *Bangladesh: Innovating for Health.* The Lancet, 2013. 382(9906), p. 1681.
360. Paykel, E.S. and Priest, R.G., *Recognition and Management of Depression in General Practice: Consensus Statement.* BMJ : British Medical Journal, 1992. 305(6863), p. 1198-1202.
361. Jin, J., *Clinicians Examine Advances and Challenges in Improving Quality of End-of-Life Care in the Icu.* JAMA, 2013. 310(23), p. 2493-2495.
362. Hayward, R.A. and Hofer, T.P., *Estimating Hospital Deaths Due to Medical Errors: Preventability Is in the Eye of the Reviewer.* JAMA, 2001. 286(4), p. 415-420.
363. Smyth, C., *Million Patients a Year Put in Hospital by Medicine Mistakes*, in *The Sunday Times*. 2017, The Times.
364. Shannon, R.P., Frndak, D., Grunden, N., Lloyd, J.C., Herbert, C., Patel, B., et al., *Using Real-Time Problem Solving to Eliminate Central Line Infections.* The Joint Commission Journal on Quality and Patient Safety, 2006. 32(9), p. 479-487.
365. Kirch, W. and Schafii, C., *Misdiagnosis at a University Hospital in 4 Medical Eras Report on 400 Cases.* Medicine, 1996. 75(1), p. 29-40.
366. Van Rensburg, H., *Health and Health Care in South Africa*. 2004, Van Schaik Publishers, Cape Town
367. Watera, C., Todd, J., Muwonge, R., Whitworth, J., Nakiyingi-Miiro, J., Brink, A., et al., *Feasibility and Effectiveness of Cotrimoxazole*

Prophylaxis for Hiv-1-Infected Adults Attending an Hiv/Aids Clinic in Uganda. JAIDS Journal of Acquired Immune Deficiency Syndromes, 2006. 42(3), p. 373-378.
368. Ho, M., Marger, M., Beart, J., Yip, I., and Shekelle, P., *Is the Quality of Diabetes Care Better in a Diabetes Clinic or in a General Medicine Clinic?* Diabetes Care, 1997. 20(4), p. 472-475.
369. Maclean, A. and Renwick, C., *Audit of Pre-Operative Starvation.* Anaesthesia, 1993. 48(2), p. 164-166.
370. Author Withheld, *Reference Withheld for Reviewing Purposes.* 2020.
371. Kale, V.P., Joshi, G.S., Gohil, P.B., and Jain, M.R., *Effect of Fasting Duration on Clinical Pathology Results in Wistar Rats.* Veterinary Clinical Pathology, 2009. 38(3), p. 361-366.
372. Grol, R., Rooijackers-Lemmers, N., Van Kaathoven, L., Wollersheim, H., and Mokkink, H., *Communication at the Interface: Do Better Referral Letters Produce Better Consultant Replies?* British Journal of General Practice, 2003. 53(488), p. 217-219.
373. Benjamin, R.M., *Medication Adherence: Helping Patients Take Their Medicines as Directed.* Public Health Reports, 2012. 127(1), p. 2-3.
374. Cutler, D.M. and Everett, W., *Thinking Outside the Pillbox—Medication Adherence as a Priority for Health Care Reform.* New England Journal of Medicine, 2010.
375. Dimatteo, M.R., Giordani, P.J., Lepper, H.S., and Croghan, T.W., *Patient Adherence and Medical Treatment Outcomes a Meta-Analysis.* Medical care, 2002, p. 794-811.
376. Ellenbecker, C.H., Frazier, S.C., and Verney, S., *Nurses' Observations and Experiences of Problems and Adverse Effects of Medication Management in Home Care.* Geriatric Nursing, 2004. 25(3), p. 164-170.
377. Mccarthy, R., *The Price You Pay for the Drug Not Taken.* Business and Health, 1998. 16(10), p. 27.
378. Burnier, M., *Medication Adherence and Persistence as the Cornerstone of Effective Antihypertensive Therapy.* American Journal of Hypertension, 2006. 19(11), p. 1190-6.
379. Bedell, S.E., Jabbour, S., Goldberg, R., Glaser, H., Gobble, S., Young-Xu, Y., et al., *Discrepancies in the Use of Medications: Their Extent and Predictors in an Outpatient Practice.* Archives of Internal Medicine, 2000. 160(14), p. 2129-2134.

380. Fund Commonwealth, *International Health Policy Survey of Sicker Adults.* Funding for Germany was provided by the German Institute for Quality and Economic Efficiency in Health Care, 2005.
381. Mulhem, E., Lick, D., Varughese, J., Barton, E., Ripley, T., and Haveman, J., *Adherence to Medications after Hospital Discharge in the Elderly.* International Journal of Family Medicine, 2013. 2013.
382. Un General Assembly, *Universal Declaration of Human Rights.* UN General Assembly, 1948.
383. Backman, G., Hunt, P., Khosla, R., Jaramillo-Strouss, C., Fikre, B.M., Rumble, C., et al., *Health Systems and the Right to Health: An Assessment of 194 Countries.* The Lancet, 2008. 372(9655), p. 2047-2085.
384. Hurst, J., *Performance Measurement and Improvement in Oecd Health Systems: Overview of Issues and Challenges.* Measuring up: improving health system performance in OECD countries, 2002.
385. Kaplan, T. and Pear, R., *New York Edition with the Headline: Unity Is Elusive as G.O.P. Presses Health Overhaul*, in *New York Times*. 2017, Arthur Ochs Sulzberger Jr., New York City. p. 1.
386. Blumenthal, D. and Hsiao, W., *Privatization and Its Discontents—the Evolving Chinese Health Care System*. 2005, Mass Medical Soc.
387. Hsiao, W.C., *The Chinese Health Care System: Lessons for Other Nations.* Social Science & Medicine, 1995. 41(8), p. 1047-1055.
388. Hall, V., Thomsen, R.W., Henriksen, O., and Lohse, N., *Diabetes in Sub Saharan Africa 1999-2011: Epidemiology and Public Health Implications. A Systematic Review.* BMC Public Health, 2011. 11(1), p. 564.
389. Leive, A. and Xu, K., *Coping with out-of-Pocket Health Payments: Empirical Evidence from 15 African Countries.* Bulletin of the World Health Organization, 2008. 86, p. 849-856C.
390. Su, T.T., Kouyaté, B., and Flessa, S., *Catastrophic Household Expenditure for Health Care in a Low-Income Society: A Study from Nouna District, Burkina Faso.* Bulletin of the World Health Organization, 2006. 84, p. 21-27.
391. Broyles, J.R. *Estimating Business Loss to a Hospital Emergency Department from Patient Reneging by Queuing-Based Regression*. in *IIE Annual Conference. Proceedings*. 2007. Institute of Industrial and Systems Engineers (IISE).

392. Balderston, D., Gonzalez, M., and López, A.M., *Encyclopedia of Contemporary Latin American and Caribbean Cultures*. 2000, Routledge, Abingdon, UK
393. Van Oosterhout, J.J., Bodasing, N., Kumwenda, J.J., Nyirenda, C., Mallewa, J., Cleary, P.R., et al., *Evaluation of Antiretroviral Therapy Results in a Resource-Poor Setting in Blantyre, Malawi.* Tropical Medicine & International Health, 2005. 10(5), p. 464-470.
394. Hartmann, D. and Sunjka, B.P., *Private Theatre Utilisation in South Africa: A Case Study.* South African Medical Journal, 2013. 103(5), p. 285 - 287.
395. Hannan, E.L., Giglio, R.J., and Sadowski, R.S. *A Simulation Analysis of a Hospital Emergency Department*. in *Proceedings of the 7th conference on Winter simulation-Volume 1*. 1974. ACM.
396. Paul, S.A., Reddy, M.C., and Deflitch, C.J., *A Systematic Review of Simulation Studies Investigating Emergency Department Overcrowding.* SIMULATION, 2010. 86(8-9), p. 559-571.
397. Ghobbar, A.A. and Friend, C.H., *Sources of Intermittent Demand for Aircraft Spare Parts within Airline Operations.* Journal of Air Transport Management, 2002. 8(4), p. 221-231.
398. Drucker, P., *The Effective Executive: The Definitive Guide to Getting the Right Things Done*. 2006, Collins, New York
399. Bauer, K., *Kpis: Avoiding the Threshold Mcguffins.* Information Management, 2005. 15(4), p. 47.
400. Muller, J.Z., *The Tyranny of Metrics*. 2018, Princeton University Press
401. Goodhart, C.a.E., *Monetary Theory and Practice: The Uk-Experience*. 1983, Macmillan International Higher Education
402. Strathern, M., *'Improving Ratings': Audit in the British University System.* European Review, 1997. 5(3), p. 305-321.
403. Supply Chain Council, *Supply Chain Operations Reference Model.* Overview of SCOR version, 2008. 5(0).
404. Krauth, E., Moonen, H., Popova, V., and Schut, M. *Performance Indicators in Logistics Service Provision and Warehouse Management–a Literature Review and Framework*. in *Euroma international conference*. 2005.
405. Keebler, J.S., Manradt, K.B., Durtsche, D.A., and Ledyard, D.M., *Keeping Score: Measuring the Business Value of Logistics in the Supply Chain*. 1999, Council for Logistics Management, Boston, Massachusetts
406. Slack, N., *The Manufacturing Advantage*. 1991, Mercury, London

407. Da Silva Moitinho, M., Das Chagas, H.X., Matos, M.C.P., De Oliveira, R.M.N., De Souza, J.F., and De Moura, V.A. *The Use of Indicators to Measure the Performance of Logistics Services in a Steel Industry*. in *PONSMeeting*. Brasilia, Brazil.
408. Fawcett, S.E. and Cooper, M.B., *Logistics Performance Measurement and Customer Success.* Industrial Marketing Management, 1998. 27(4), p. 341-357.
409. Coyle, J.J., Langley, C.J., Novack, R.A., and Gibson, B., *Supply Chain Management: A Logistics Perspective*. 2016, Nelson Education, Toronto, Canada
410. Pycraft, M., Singh, H., Phihlela, K., Slack, N., Chambers, S., and Johnston, R., *Operations Management: Global and Southern African Perspectives.* Cape Town: Pearson Education, 2010.
411. Heizer, J. and Render, B., *Principles of Operations Management*. 2011, Pearson, London
412. Huan, S.H., Sheoran, S.K., and Wang, G., *A Review and Analysis of Supply Chain Operations Reference (Scor) Model.* Supply Chain Management: An International Journal, 2004. 9(1), p. 23-29.
413. World Health Organization, *Annex 5 Who Good Distribution Practices for Pharmaceutical Products.* 2010.
414. Monden, Y. and Lee, J., *How a Japanese Auto Maker Reduces Costs.* Management Accounting (USA), 1993. 75(2), p. 22-27.
415. Rushton, A., Croucher, P., and Baker, P., *The Handbook of Logistics and Distribution Management*. 5th Ed ed. 2019, KoganPage, London
416. Aghezzaf, E., *Capacity Planning and Warehouse Location in Supply Chains with Uncertain Demands.* Journal of the Operational Research Society, 2005. 56(4), p. 453-462.
417. Jain, C.L., *Benchmarking Forecasting Models.* The Journal of Business Forecasting, 2006. 25(4), p. 14.
418. Box, G.E., Jenkins, G.M., Reinsel, G.C., and Ljung, G.M., *Time Series Analysis: Forecasting and Control*. 2015, John Wiley & Sons, Hoboken, New Jersey
419. Hyndman, R.J. and Koehler, A.B., *Another Look at Measures of Forecast Accuracy.* International Journal of Forecasting, 2006. 22(4), p. 679-688.
420. Wheelwright, S. and Makridakis, S., *Forecasting Methods for Managers.* 1974.

421. Fumi, A., Pepe, A., Scarabotti, L., and Schiraldi, M.M., *Fourier Analysis for Demand Forecasting in a Fashion Company.* International Journal of Engineering Business Management, 2013. 5(Godište 2013), p. 5-30.
422. Khalil Zadeh, N., Sepehri, M.M., and Farvaresh, H., *Intelligent Sales Prediction for Pharmaceutical Distribution Companies: A Data Mining Based Approach.* Mathematical Problems in Engineering, 2014. 2014.
423. Albright, S.C., *Vba for Modelers: Developing Decision Support Systems with Microsoft Excel*. 4th ed. 2011, Cengage Learning, Inc;
424. Chassaing, P. and Gerin, L., *Efficient Estimation of the Cardinality of Large Data Sets.* arXiv Preprint Math/0701347, 2007.
425. Rother, M. and Shook, J., *Learning to See: Value Stream Mapping to Add Value and Eliminate Muda*. 2003, Lean Enterprise Institute
426. Richardson, T. and Richardson, E., *The Toyota Engagement Equation: How to Understand and Implement Continuous Improvement Thinking in Any Organization*. 1st edition ed. 2017, McGraw-Hill Education, New York City
427. World Health Organization, *Everybody's Business--Strengthening Health Systems to Improve Health Outcomes: Who's Framework for Action.* 2007.
428. Lansing, J.S., *Complex Adaptive Systems.* Annual Review of Anthropology, 2003. 32(1), p. 183-204.
429. Agrawal, R. and Srikant, R. *Mining Sequential Patterns*. in *International Conference on Data Engineering*. 1995.
430. Mabroukeh, N.R. and Ezeife, C.I., *A Taxonomy of Sequential Pattern Mining Algorithms.* ACM Computing Surveys (CSUR), 2010. 43(1), p. 3.
431. Green, L.W., *Public Health Asks of Systems Science: To Advance Our Evidence-Based Practice, Can You Help Us Get More Practice-Based Evidence?* American Journal of Public Health, 2006. 96(3), p. 406-409.
432. Campbell, S.M., Roland, M.O., and Buetow, S.A., *Defining Quality of Care.* Social Science & Medicine, 2000. 51(11), p. 1611-1625.
433. Jun, J.B., Jacobson, S.H., and Swisher, J.R., *Application of Discrete-Event Simulation in Health Care Clinics: A Survey.* The journal of the Operational Research Society, 199. 50(2), p. 109-123.
434. Jacobson, N., Butterill, D., and Goering, P., *Development of a Framework for Knowledge Translation: Understanding User*

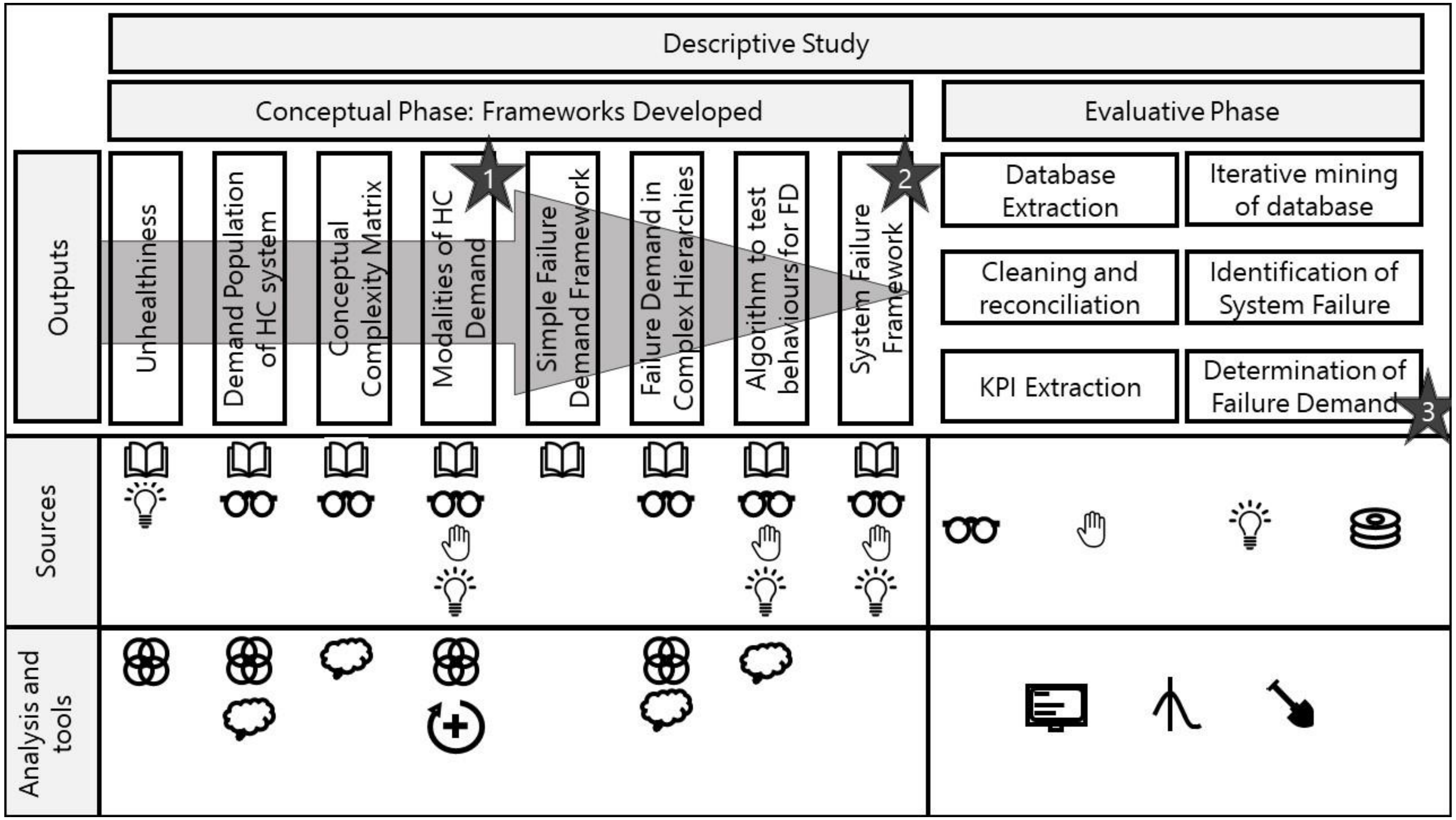
Context. Journal of Health Services Research & Policy, 2003. 8(2), p. 94-99.
435. Reid, P.P., Compton, W.D., Grossman, J.H., and Fanjiang, G., *Building a Better Delivery System: A New Engineering/Health Care Partnership*. 2005, National Academies Press, Washington D.C.
436. Andersen, B. and Fagerhaug, T., *Root Cause Analysis: Simplified Tools and Techniques*. 2006, ASQ Quality Press, Milwaukee, Wisconsin

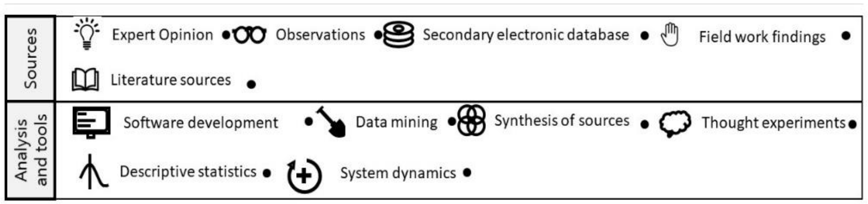
9 - 1. Appendix 1: Structure of the thesis

Central Research Question: What impact does failure demand have on the gap between supply and demand in health systems?

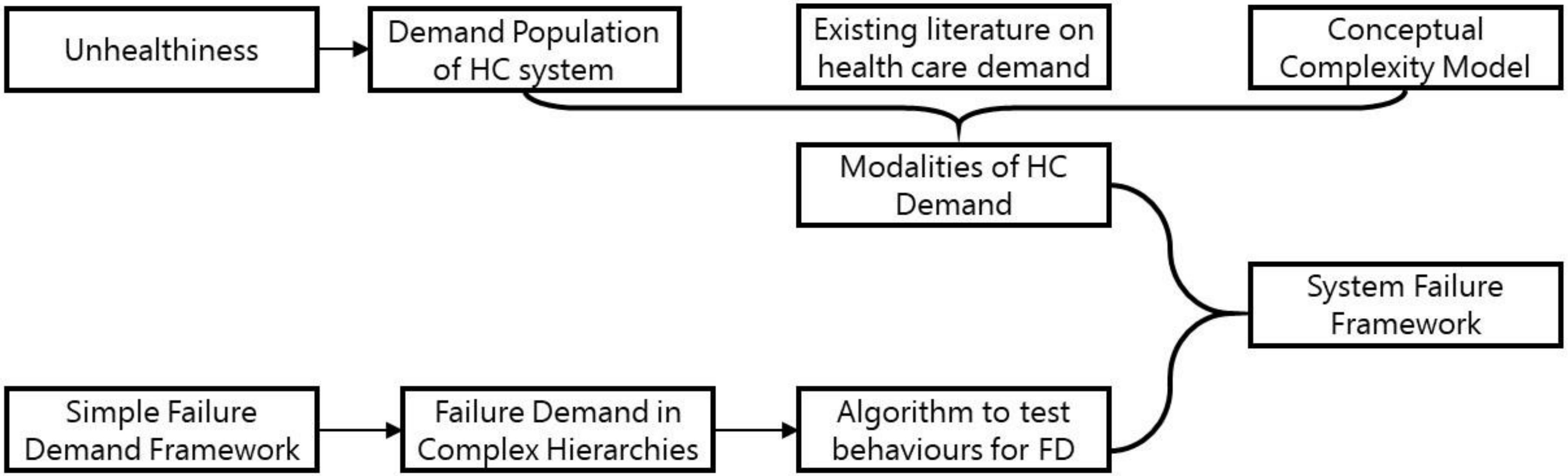
Objectives	**Papers**	**Sub Question:**	**Chapter**
Create a systemic demand modality model	"Demand modalities in health care"	What is the natural demand population of the health system? Which demand modalities exist in this population? Can demand be described in system dynamic terms? Can a complexity based demand model be proposed?	4
Create a framework to identify and categorise failure demand in complex hierarchical systems	"Understanding Failure demand in Health care: A Mental Model for Demand Management"	Can a failure demand be elaborated on for a complex setting? Can a generic algorhithm be designed to assess failure demand in hierarchical settings? Can a specific mental model be constructed to describe failure demand in a health setting?	5
Measure an example of the prevalence of failure demand in a health setting	"Measuring Failure Demand in a National Pharmaceutical Supply Chain"	What proportion of system failure exists in a health care setting? What proportion of demand is a result of failure demand in a health care setting?	6

9 - 2. Appendix 2: Overall study structure

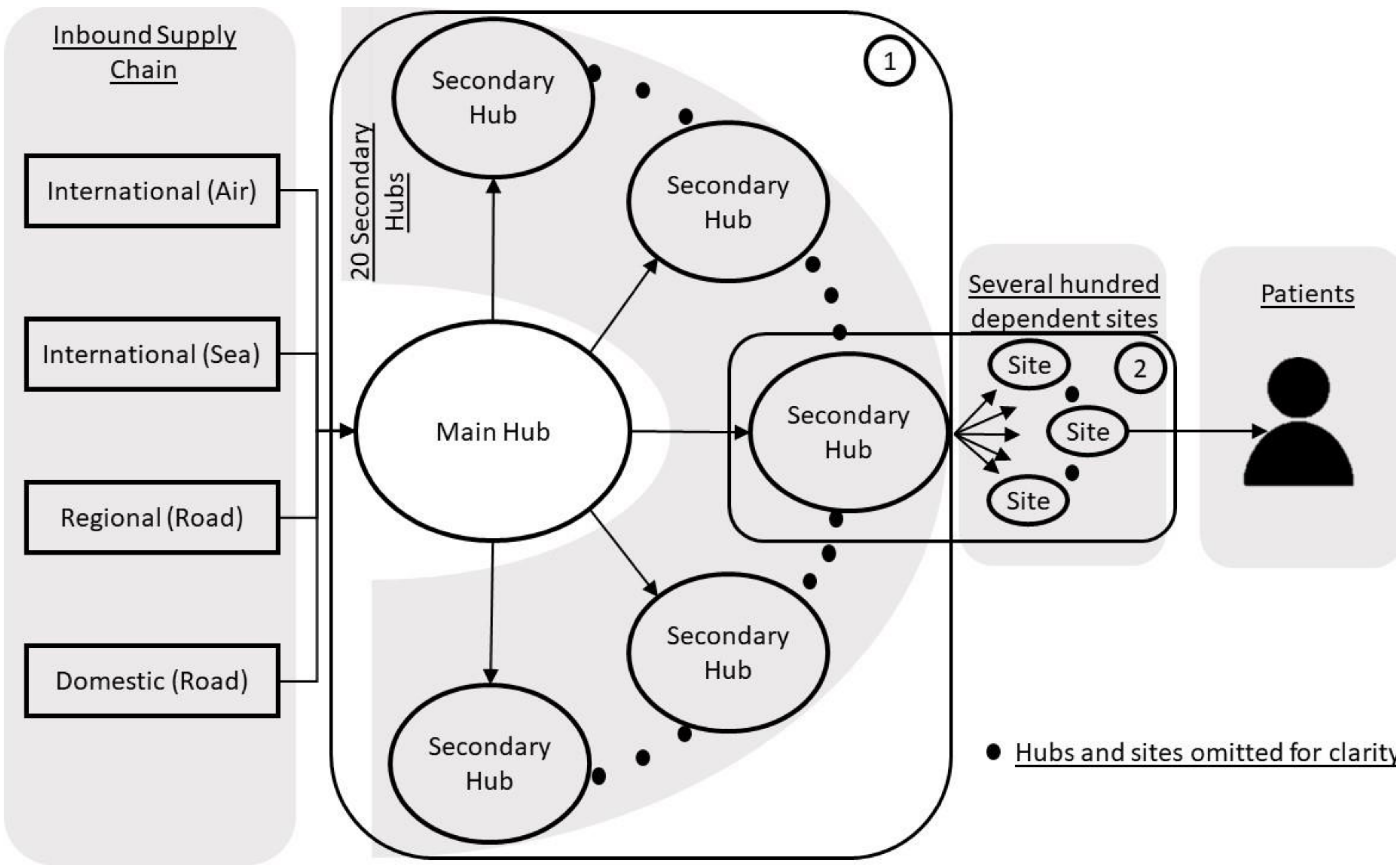




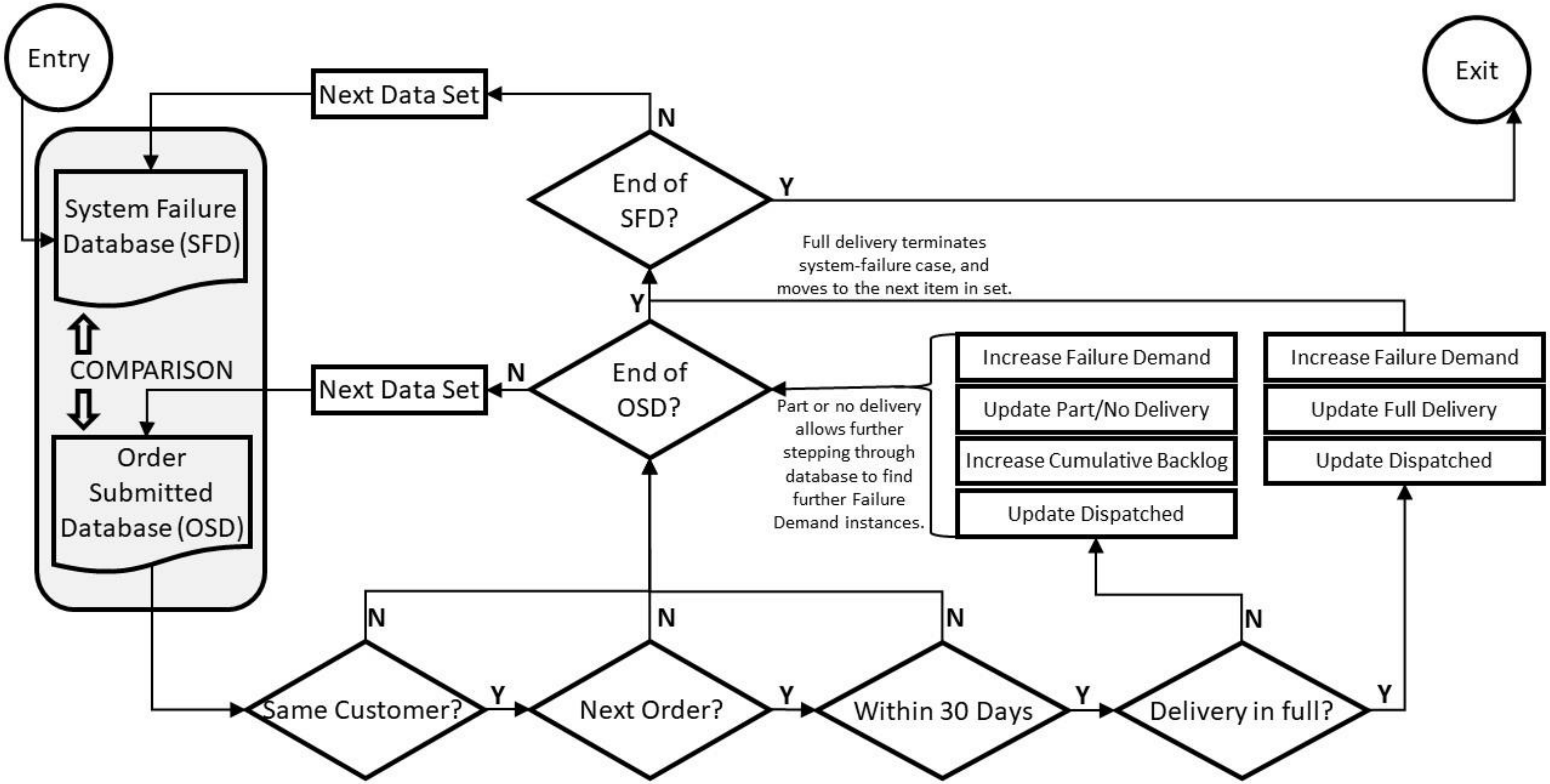
9 - 3. Appendix 3: Conceptual phase



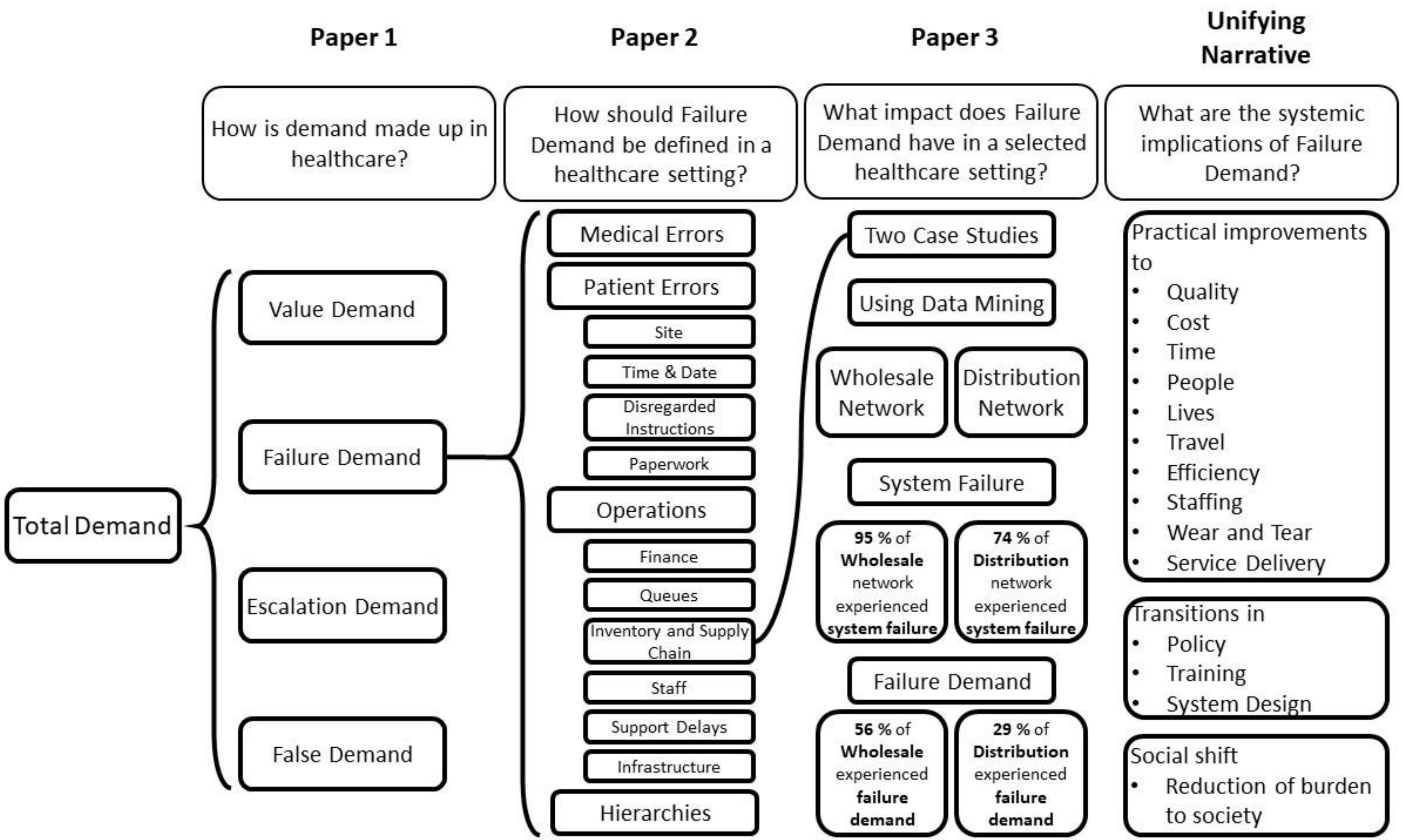
9 - 4. Appendix 4: Pharmaceutical supply network structure



9 - 5. Appendix 5: Analytical algorithm



9 - 6. Appendix 6: Flow of the thesis



9 - 7. Appendix 7: Performance metrics in supply networks

Dimensions for competitive advantage	SCOR Model	Keeping score	Metrics
Quality	-	Quality	Customer Satisfaction, Processing Accuracy, Order Fulfilment (on time, in full, accurate selection, damage free) Inventory Accuracy, Forecast Accuracy, Planning Accuracy
Cost	Cost	Cost	Inventory turns, sales outstanding, cost to serve, cash to cash cycle, total costs (cost of goods, transportation, carrying cost, MH costs, capacity costs, lost opportunity costs. Cost of goods sold,
Speed	Responsiveness	Time	On time delivery, order cycle time, cycle time variability, response time, forecasting cycle time, order lead times
Flexibility	Flexibility	-	Supply-chain response times, production flexibility.
Dependability	Reliability		Delivery Performance, Fill Rates, order fulfilment.
-	Asset Management Efficiency	-	Cash to cash cycle times, Stock Reach, Asset turns.
-	-	Uncategorised.	Minimum order Quantities, Change order Timing

... ooo 000 OOO END OOO 000 ooo ...