HYDROTHERAPY IN THE MANAGEMENT OF CHRONIC OBSTRUCTIVE PULMONARY DISEASE: A QUALITATIVE SYSTEMATIC REVIEW

Dorothy A Shead

A dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Masters of Science (Physiotherapy) Johannesburg, 2010.
DECLARATION

I, Dorothy Shead, declare that this dissertation is my own work. It is being submitted for the degree of Masters of Science (Physiotherapy) in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other University.

[Signature]

......Tenth......day of September......, 2010
ABSTRACT

Background:
Chronic obstructive pulmonary disease (COPD), characterised by progressive airflow limitation which is not fully reversible and associated with pulmonary and systemic inflammatory processes, is largely associated with smoking and is classified as a disease of lifestyle. Other factors related to the incidence of COPD are passive smoking, the inhalation of gases from biomass fuels and the genetic absence of the protease alpha_{1}-antitrypsin in some people. COPD is found in all sectors of society and is not dependent on level of income or on the global location of a population. Compounding the incidence of COPD in Africa is the human immunodeficiency virus/ acquired immunodeficiency syndrome (HIV/AIDS) that predisposes patients to the development of COPD. Hydrotherapy has been used since ancient times as a preventative and/or therapeutic form of treatment. Physiotherapists have used hydrotherapy alone or as a useful adjunct to other treatment options for years. Today aquatic therapy is practised in many centres where a multi-disciplinary approach to this form of treatment is offered. There are a number of methodologies of application of the therapy. Among these are Ai Chi, Halliwick, Watsu and Bad Ragaz. The treatment takes place in thermoneutral water between 29 degrees centigrade (°C) and 34°C where use is made of the buoyancy and/or resistance created by the aquatic medium. Musculoskeletal conditions, including pre-operative total hip and knee replacements; osteoarthritis and rheumatoid arthritis can be treated with this modality. Hydrotherapy affords athletes a non-weightbearing rehabilitative environment to enable a quicker recovery and also enables cardiovascular reconditioning to commence sooner than would be afforded by a land-based rehabilitation programme. Many chronic conditions, including cardiac failure, stroke and metabolic disorders have benefited from a hydrotherapy intervention. COPD has been shown to have an inflammatory component. Exercise on land has been shown to have an anti-inflammatory effect in healthy individuals but increased levels of tumour necrosis factor–alpha (TNF-alpha), known to be associated with inflammation, has resulted from moderate intensity land-based exercise in patients with COPD.

Aim and objectives of the research:
The aim of this study is to systematically review the evidence related to the role of hydrotherapy in the management of patients with COPD. The research objectives of this study are: to determine, from the literature, whether water temperature and depth of immersion influences cardiopulmonary function in patients with COPD; to determine, from the literature, whether hydrotherapy increases endurance, function and quality of life (QoL) in patients with COPD and to determine, from the literature, whether the anti-inflammatory effects of water-based exercise in patients with COPD have been documented.
Method:
The study design is a qualitative systematic review. A search was made for relevant journal articles in the PUBMED, SCOPUS, CINAHL, MEDLINE, SPORTSDiscus and Cochrane review databases. Google scholar was perused in order to find any grey literature pertaining to the population under review. The review of the literature was from 1996 until 2009. This protracted period of 13 years was needed to insure that all leading articles on the subject under review were included in the review. There were too few articles available from 2000 onwards to produce a valid review of the topic. The words and/or phrases used in the search were: hydrotherapy, Halliwick, Ai Chi, Watsu, Bad Ragaz, chronic obstructive pulmonary disease (COPD), pulmonary rehabilitation, emphysema, chronic bronchitis and the anti-inflammatory effects of exercise. Articles in the form of systematic reviews, randomised controlled trials, clinical trials and case studies pertaining to adult patients with non-acute, non-terminal COPD and the use of hydrotherapy were included in this review. No articles based on animal studies were included. No articles on Kneipp therapy were included as hydrotherapy in this format does not include the immersion of the patient in water. The LOW (Lewis, Olds and Williams, 2007) critical appraisal tool was used to evaluate the articles included in the systematic review.

Results:
Thirty-seven studies were sourced and nine studies were included in the qualitative systematic review. A total of 35 outcomes measures were reported within the nine included studies. They were of a varied nature and therefore were grouped into the following categories: cardiovascular; pulmonary; muscle strength/weakness; endurance and power, QoL and activities of daily living (ADL). Beneficial results were found in relation to heart rate (HR) with a regime of upper limb exercises performed in water. Following an upper body and upper limb 15 minute land-based exercise programme and subsequent 10-15 minute rest period on land, systolic blood pressure (SBP) was decreased by 14mmHg and diastolic blood pressure (DBP) by 6 mmHg (compared to resting land values) when patients with COPD were immersed in 32°C water. Ejection fraction (EF) improved significantly at the end of a two month breathing exercise programme in water. A water-based intervention period of 120 minutes (20 minutes, six days/week for two months) decreased left ventricular end–diastolic (LVd) and left ventricular end-systolic (LVs) dimensions (p<0.01). Exhaling into water, during an aquatic breathing exercise regimen lasting 30 minutes/day, six days/week for two months, was demonstrated to significantly increase percentage predicted forced expiratory volume in one second (FEV1%). Arterial concentration of carbon dioxide (PaCO2) levels were decreased significantly due to the breathing exercise with exhalation into water regimens and during the breathing out into water intervention for 20 minute /day for six days the arterial concentration of oxygen (PaO2) levels were increased. Peak flow (PF) was improved in all the breathing programmes where the patient exhaled into water. Improvement in respiratory rate (RR) and oxygen saturation (SaO2) were seen, in patients with COPD, who
performed weight-resisted upper limb exercises in water once a week with a twice weekly pulmonary rehabilitation programme (PRP) on land as opposed to a land-based PRP three times per week. Two of the included studies recorded improvement in the incremental shuttle walk test (ISWT) following exercise in water and one noted a greater improvement in the endurance shuttle walk test (ESWT) than in the ISWT after hydrotherapy. Maximal dynamic flexion showed marked improvement after an aquatic programme. Physical and cardiopulmonary improvements, including reduced levels of dyspnoea in some instances, were reported and these were linked to increased levels of physical conditioning, better QoL and improved ability to undertake ADL in the patients with COPD.

**Conclusion:**
Breathing exercises, where the patients exhale into the water, appear to have a beneficial effect on pulmonary outcomes particularly when the programme is of a duration of 120 minutes per week or more and the exercise is performed on six days/week. The physical exercise hydrotherapy programmes address some of the muscular weaknesses resulting from the systemic effects of COPD. Both the cardiopulmonary benefits and physical benefits seem to result in a general improvement in the QoL of the patients and their greater ability to perform ADL. Social interaction and psychological well-being seem to be factors related to increased compliance in hydrotherapy exercise programmes when compared to compliance in land-based programmes. No information was retrieved from the included studies relating to the anti-inflammatory effects of hydrotherapy exercise programmes. No randomised controlled trials were sourced on the subject under review. The overall evidence was of variable quality, with three studies above average, two average and four below average, according to the LOW critical appraisal tool. From the results obtained in the review it became apparent that there is an urgent need for a number of randomised controlled trials to investigate the role of breathing exercises in combination with physical exercise programmes of hydrotherapy in the management of COPD so that this form of therapy can be utilised to its full capacity.
ACKNOWLEDGEMENTS

Firstly, my sincere thanks have to go to my supervisor Dr Heleen van Asweegen who has guided me, advised me and taught me so much during the process of the completion of this dissertation. She has enriched this experience with her vast knowledge in the field of cardiopulmonary physiotherapy. When the going was tough she encouraged and when things were going well she offered congratulations. The fact that she is meticulous in her work and an excellent researcher has been an incentive for me to do my best to produce a dissertation of which she and I can be proud.

I also want to sincerely thank Professor Aimee Stewart. When I decided to embark on the MSc Physiotherapy path it was with some fear and trepidation as it was some length of time since I had embraced academia. Professor Stewart was most encouraging and supportive. She made me believe that I was capable of completing this degree and I hope by completing this dissertation I will have rewarded her faith in me.

After meeting with the Postgraduate Committee it was recommended that I talk to Ms Patricia McInerney as she was an authority on the systematic review process. I met with Patricia and she was most helpful in guiding me as to the format for my systematic review. I thank Patricia for guiding me and helping me to formulate the structure of this qualitative systematic review.

I wish to thank Lucy Lewis, of the School of Health Sciences (Physiotherapy), University of South Australia in Adelaide for granting me permission to utilise the LOW critical appraisal tool in this dissertation and for her advice on the method of utilisation of the tool. Her words of encouragement, regarding my achieving my goal of gaining an MSc in Physiotherapy are much appreciated.

My thanks have to be extended to the entire Physiotherapy department of the University of the Witwatersrand, Johannesburg, South Africa as they have all been very supportive and helpful offering ideas, advice and generally encouraging my endeavours.

On a personal note I wish to thank my very special daughter Michelle for her belief in my abilities, her encouragement and her love.

I would also like to thank all my long-suffering friends who have had to listen to my drawn-out deliberations, moans and occasional elated moments during the creation of this dissertation. They are all looking forward to a dissertation-free time.

Finally I would like to dedicate this dissertation to the memory of my late son Gareth.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declaration</td>
<td>ii</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>vi</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>v</td>
</tr>
<tr>
<td>List of Tables</td>
<td>ix</td>
</tr>
<tr>
<td>List of Figures</td>
<td>x</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>xi</td>
</tr>
</tbody>
</table>

## 1. Chapter 1: Introduction

1.1 Background                                                          | 1    |
1.2 Statement of Problem and Justification for Research                | 5    |
1.3 Research Question                                                  | 5    |
1.4 Significance of Research                                           | 5    |
1.5 Research Aim                                                       | 6    |
1.6 Research Objectives                                                | 6    |
1.7 Type of Study                                                      | 6    |

## 2. Chapter 2: Literature Review

2.1 Factors Predisposing to the Development of COPD                    | 7    |
2.1.1 The Role of Cellular Mediators in the Development of COPD        | 10   |
2.2 Classification of Severity of COPD                                  | 12   |
2.3 The Influence of COPD on Endurance and Function in Activities of Daily Living | 13   |
2.4 The Influence of COPD on Quality of Life                           | 16   |
2.5 The Medical Management of Patients with COPD                       | 19   |
2.6 Rehabilitation Interventions for COPD                              | 20   |
2.6.1 Pulmonary Rehabilitation for Patients with COPD                  | 21   |
2.6.2 Inspiratory Muscle Training as a Component of Pulmonary Rehabilitation | 23   |
2.6.3 Home-Based versus Hospital-Based Rehabilitation                  | 24   |
2.7 Hydrotherapy                                                       | 25   |
2.7.1 Method of Hydrotherapy Application                               | 25   |
2.7.2 The Effect of Hydrostatic pressure on the Cardiovascular and Respiratory Systems in Healthy Individuals | 26   |
2.7.3 The Effect of Different Levels of Immersion on the Cardiovascular and Respiratory Systems in Healthy Individuals | 28   |
2.7.4 The Effect of Different Water Temperatures on the Cardiovascular and Respiratory Systems in Healthy Individuals | 28   |
2.7.5 Submaximal and Maximal Workloads in Water for Healthy Individuals | 30   |
2.7.6 The Role of Hydrotherapy in the Management of Chronic Diseases   | 32   |
2.7.7 The Pathophysiological Effect of Cold Water Therapy              | 34   |
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.</td>
<td>Chapter 3: Methodology</td>
<td>35</td>
</tr>
<tr>
<td>3.1</td>
<td>Study Design</td>
<td>35</td>
</tr>
<tr>
<td>3.2</td>
<td>Research Hypotheses</td>
<td>35</td>
</tr>
<tr>
<td>3.3</td>
<td>Search Strategy</td>
<td>35</td>
</tr>
<tr>
<td>3.4</td>
<td>Inclusion Criteria</td>
<td>36</td>
</tr>
<tr>
<td>3.5</td>
<td>Exclusion Criteria</td>
<td>36</td>
</tr>
<tr>
<td>3.6</td>
<td>Data Collection</td>
<td>36</td>
</tr>
<tr>
<td>3.7</td>
<td>Data Analysis</td>
<td>36</td>
</tr>
<tr>
<td>3.8</td>
<td>Ethical Considerations</td>
<td>36</td>
</tr>
<tr>
<td>4.</td>
<td>Chapter 4: Results</td>
<td>38</td>
</tr>
<tr>
<td>4.1</td>
<td>Identification and Selection of Studies</td>
<td>38</td>
</tr>
<tr>
<td>4.2</td>
<td>Description of the Studies</td>
<td>40</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Volume of Evidence According to Lloyd-Smith (1997) Hierarchy</td>
<td>40</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Generalisability of evidence</td>
<td>45</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Participants</td>
<td>45</td>
</tr>
<tr>
<td>4.2.4</td>
<td>Intervention</td>
<td>46</td>
</tr>
<tr>
<td>4.2.5</td>
<td>Outcome Measures Assessed</td>
<td>47</td>
</tr>
<tr>
<td>4.2.5.1</td>
<td>The Cardiovascular Outcomes</td>
<td>47</td>
</tr>
<tr>
<td>4.2.5.2</td>
<td>The Pulmonary Outcomes</td>
<td>49</td>
</tr>
<tr>
<td>4.2.5.3</td>
<td>Muscle Strength/Weakness, Endurance and Power</td>
<td>51</td>
</tr>
<tr>
<td>4.2.5.4</td>
<td>Quality of Life and Activities of Daily Living Outcomes</td>
<td>52</td>
</tr>
<tr>
<td>5.</td>
<td>Chapter 5: Discussion</td>
<td>54</td>
</tr>
<tr>
<td>6.</td>
<td>Chapter 6: Conclusions and recommendations</td>
<td>62</td>
</tr>
<tr>
<td>7.</td>
<td>References</td>
<td>63</td>
</tr>
<tr>
<td>Appendix A</td>
<td>Recommendations for pulmonary rehabilitation according to ACCP/AACVPR (2007)</td>
<td>79</td>
</tr>
<tr>
<td>Appendix B</td>
<td>Human Research Ethics Committee (Medical) clearance certificate</td>
<td>80</td>
</tr>
<tr>
<td>Appendix C</td>
<td>Change of title of research approval certificate</td>
<td>81</td>
</tr>
<tr>
<td>Appendix D</td>
<td>LOW critical appraisal scores (DS and HvA)</td>
<td>82</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table 4.1</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The allocation of Studies According to Search Engines and Keywords of Phrases</td>
<td>39</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4.2</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hierarchy of Evidence Ranking System (Lloyd-Smith, 1997)</td>
<td>41</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4.3</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methodological Quality of the Included Studies (n=9) Based on the LOW Critical Appraisal Tool</td>
<td>43</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure 4.1</th>
<th>Study Selection Process</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 4.2</td>
<td>The LOW (Lewis, Olds and Williams, 2007) critical appraisal tool</td>
<td>42</td>
</tr>
</tbody>
</table>
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AACVPR</td>
<td>American Association of Cardiovascular and Pulmonary Rehabilitation</td>
</tr>
<tr>
<td>ACCP</td>
<td>American College of Chest Physicians</td>
</tr>
<tr>
<td>ACSM</td>
<td>American College of Sports Medicine</td>
</tr>
<tr>
<td>ADL</td>
<td>Activities of Daily Living</td>
</tr>
<tr>
<td>AIDS</td>
<td>Acquired Immunodeficiency Syndrome</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Data/Variance</td>
</tr>
<tr>
<td>ATS</td>
<td>American Thoracic Society</td>
</tr>
<tr>
<td>BAL</td>
<td>Bronchoalveolar Lavage</td>
</tr>
<tr>
<td>BASDEC</td>
<td>Brief assessment of Depression Cards</td>
</tr>
<tr>
<td>BE</td>
<td>Base Excess</td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
</tr>
<tr>
<td>BODE</td>
<td>Body Mass Index/Airflow Obstruction/ Dyspnoea and Exercise Capacity Index.</td>
</tr>
<tr>
<td>BOLD</td>
<td>Burden of Obstructive Lung Disease</td>
</tr>
<tr>
<td>BPQ</td>
<td>Breathing Problems Questionnaire</td>
</tr>
<tr>
<td>CARS</td>
<td>COPD Activity Rating Scale</td>
</tr>
<tr>
<td>CBV</td>
<td>Central Blood Volume</td>
</tr>
<tr>
<td>CCQ</td>
<td>Clinical Chronic Obstructive Pulmonary Disease Questionnaire</td>
</tr>
<tr>
<td>CET</td>
<td>Cycle Ergometry Training</td>
</tr>
<tr>
<td>CO</td>
<td>Cardiac Output</td>
</tr>
<tr>
<td>COPD</td>
<td>Chronic Obstructive Pulmonary Disease</td>
</tr>
<tr>
<td>CRQ</td>
<td>Chronic Respiratory Questionnaire</td>
</tr>
<tr>
<td>CRQ-SR</td>
<td>Chronic Respiratory Questionnaire Self-Reported.</td>
</tr>
<tr>
<td>CSA</td>
<td>Cross-Sectional Area</td>
</tr>
<tr>
<td>CSM</td>
<td>Cigarette Smoke Medium</td>
</tr>
<tr>
<td>CV</td>
<td>Closing Volume</td>
</tr>
<tr>
<td>CVP</td>
<td>Central Venous Pressure</td>
</tr>
<tr>
<td>°C</td>
<td>Degrees Centigrade</td>
</tr>
<tr>
<td>DBP</td>
<td>Diastolic Blood Pressure</td>
</tr>
<tr>
<td>DWR</td>
<td>Deep Water Running</td>
</tr>
<tr>
<td>EF</td>
<td>Ejection Fraction</td>
</tr>
<tr>
<td>EMT</td>
<td>Expiratory Muscle Training</td>
</tr>
<tr>
<td>ERS</td>
<td>European Respiratory Society</td>
</tr>
<tr>
<td>ES</td>
<td>Effect Size</td>
</tr>
<tr>
<td>ESWT</td>
<td>Endurance Shuttle Walking Test</td>
</tr>
<tr>
<td>FEV₁</td>
<td>Forced Expiratory Volume In One Second</td>
</tr>
<tr>
<td>FEV₁%</td>
<td>Forced Expiratory Volume in One Second (percentage predicted)</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>FRC</td>
<td>Functional Residual Capacity</td>
</tr>
<tr>
<td>FVC</td>
<td>Forced Vital Capacity</td>
</tr>
<tr>
<td>GER</td>
<td>General Exercise Re-Conditioning</td>
</tr>
<tr>
<td>GET</td>
<td>Gas Exchange Threshold</td>
</tr>
<tr>
<td>GOLD</td>
<td>Global Initiative for Chronic Obstructive Lung Disease</td>
</tr>
<tr>
<td>HAD</td>
<td>Hospital Anxiety and Depression Scale</td>
</tr>
<tr>
<td>HIV</td>
<td>Human Immunodeficiency Virus</td>
</tr>
<tr>
<td>HR</td>
<td>Heart Rate</td>
</tr>
<tr>
<td>HR&lt;sub&gt;peak&lt;/sub&gt;</td>
<td>Peak Heart Rate</td>
</tr>
<tr>
<td>HRCT</td>
<td>High-Resolution Computed Tomography</td>
</tr>
<tr>
<td>HRQoL</td>
<td>Health Related Quality Of Life</td>
</tr>
<tr>
<td>ICS</td>
<td>Inhaled Corticosteroids</td>
</tr>
<tr>
<td>IFN-γ</td>
<td>Interferon-gamma</td>
</tr>
<tr>
<td>IL5</td>
<td>Interleukin-5</td>
</tr>
<tr>
<td>IL8</td>
<td>Interleukin-8</td>
</tr>
<tr>
<td>IMT</td>
<td>Inspiratory Muscle Training</td>
</tr>
<tr>
<td>IRV</td>
<td>Inspiratory Reserve Volume</td>
</tr>
<tr>
<td>ISWT</td>
<td>Incremental Shuttle Walk Test</td>
</tr>
<tr>
<td>LABA</td>
<td>Long-Acting β&lt;sub&gt;2&lt;/sub&gt;-Antagonists</td>
</tr>
<tr>
<td>LBP</td>
<td>Low Back Pain</td>
</tr>
<tr>
<td>LCADL</td>
<td>London Chest Activity of Daily Living scale</td>
</tr>
<tr>
<td>LT</td>
<td>Lactate Threshold</td>
</tr>
<tr>
<td>LTB&lt;sub&gt;4&lt;/sub&gt;</td>
<td>Leukotriene B&lt;sub&gt;4&lt;/sub&gt;</td>
</tr>
<tr>
<td>LOW</td>
<td>Lewis, Williams and Olds</td>
</tr>
<tr>
<td>LV&lt;sub&gt;d&lt;/sub&gt;</td>
<td>Left-Ventricular End-Diastolic</td>
</tr>
<tr>
<td>LV&lt;sub&gt;s&lt;/sub&gt;</td>
<td>Left–Ventricular End-Systolic</td>
</tr>
<tr>
<td>MACL</td>
<td>Mood-Adjective Check List</td>
</tr>
<tr>
<td>MCID</td>
<td>Minimally Clinically Important Difference</td>
</tr>
<tr>
<td>MCS</td>
<td>Mental Component Score</td>
</tr>
<tr>
<td>MRC</td>
<td>Medical Research Council</td>
</tr>
<tr>
<td>6MWD</td>
<td>Six-Minute Walk Distance</td>
</tr>
<tr>
<td>6MST</td>
<td>Six-Minute Stepper Test</td>
</tr>
<tr>
<td>NCD</td>
<td>Non-Communicable Diseases</td>
</tr>
<tr>
<td>NEADL</td>
<td>Nottingham Extended Activity of Daily Living</td>
</tr>
<tr>
<td>NHLBI</td>
<td>National Heart, Lung and Blood Institute</td>
</tr>
<tr>
<td>NK-cells</td>
<td>Natural Killer-cells</td>
</tr>
<tr>
<td>NNT</td>
<td>Numbers needed to be treated</td>
</tr>
<tr>
<td>NYHA</td>
<td>New York Heart Association</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>p value</td>
<td>Power Calculation</td>
</tr>
<tr>
<td>PaCO₂</td>
<td>Partial Pressure of Carbon Dioxide in Arterial Blood</td>
</tr>
<tr>
<td>PaO₂</td>
<td>Partial Pressure of Oxygen in Arterial Blood</td>
</tr>
<tr>
<td>PCS</td>
<td>Physical Component Score</td>
</tr>
<tr>
<td>PEF</td>
<td>Peak Expiratory Flow</td>
</tr>
<tr>
<td>PEmax</td>
<td>Maximum Expiratory Pressure</td>
</tr>
<tr>
<td>PF</td>
<td>Peak Flow</td>
</tr>
<tr>
<td>PFSDQ</td>
<td>Pulmonary Functional Status and Dyspnoea Questionnaire</td>
</tr>
<tr>
<td>pH</td>
<td>Acid/Alkaline Level</td>
</tr>
<tr>
<td>PR</td>
<td>Pulmonary Rehabilitation</td>
</tr>
<tr>
<td>PRP</td>
<td>Pulmonary Rehabilitation Programme</td>
</tr>
<tr>
<td>QLI</td>
<td>Ferrans and Powers Quality of Life Index-Pulmonary Version III</td>
</tr>
<tr>
<td>QoL</td>
<td>Quality of Life</td>
</tr>
<tr>
<td>QWB</td>
<td>Quality of Well-Being</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised Controlled Trial</td>
</tr>
<tr>
<td>RER</td>
<td>Respiratory Exchange Ratio</td>
</tr>
<tr>
<td>ROM</td>
<td>Range of Movement</td>
</tr>
<tr>
<td>RPE</td>
<td>Rate of Perceived Exertion</td>
</tr>
<tr>
<td>RPP</td>
<td>Rate-Pressure Product</td>
</tr>
<tr>
<td>RR</td>
<td>Respiratory Rate</td>
</tr>
<tr>
<td>RV</td>
<td>Residual Volume</td>
</tr>
<tr>
<td>SADHS</td>
<td>South African Demographic and Health Survey</td>
</tr>
<tr>
<td>SAMRC</td>
<td>South African Medical Research Council</td>
</tr>
<tr>
<td>SaO₂</td>
<td>Oxygen Saturation</td>
</tr>
<tr>
<td>SATS</td>
<td>South African Thoracic Society</td>
</tr>
<tr>
<td>SBP</td>
<td>Systolic Blood Pressure</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>SF36</td>
<td>Short-Form 36 Item Questionnaire</td>
</tr>
<tr>
<td>SGRQ</td>
<td>St George's Respiratory Questionnaire</td>
</tr>
<tr>
<td>SIDS</td>
<td>Sudden Infant Death Syndrome</td>
</tr>
<tr>
<td>SIP</td>
<td>Sickness Impact Profile</td>
</tr>
<tr>
<td>SV</td>
<td>Stroke Volume</td>
</tr>
<tr>
<td>SWR</td>
<td>Shallow Water Running</td>
</tr>
<tr>
<td>TMR</td>
<td>Treadmill Running</td>
</tr>
<tr>
<td>TNF-α</td>
<td>Tumour Necrosis Factor alpha</td>
</tr>
<tr>
<td>TORCH</td>
<td>Towards a Revolution in COPD Health</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>USA</td>
<td>United States of America</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>(V_{25})</td>
<td>Maximal Expiratory Flow At 25 Percent</td>
</tr>
<tr>
<td>(V_{50})</td>
<td>Maximal Expiratory Flow At 50 Percent</td>
</tr>
<tr>
<td>VAS</td>
<td>Visual Analogue Scale</td>
</tr>
<tr>
<td>VC</td>
<td>Vital Capacity</td>
</tr>
<tr>
<td>(VO_2)</td>
<td>Oxygen Consumption/Uptake</td>
</tr>
<tr>
<td>(VO_2) max</td>
<td>Maximum Oxygen Consumption</td>
</tr>
<tr>
<td>(VO_2) peak</td>
<td>Peak Oxygen Consumption</td>
</tr>
<tr>
<td>(V_T)</td>
<td>Tidal Volume</td>
</tr>
<tr>
<td>(V_{T\text{ ABD}})</td>
<td>Tidal Volume (abdominal component)</td>
</tr>
<tr>
<td>(V_{T\text{ RC}})</td>
<td>Tidal Volume (rib cage component)</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
</tbody>
</table>
CHAPTER 1

1. INTRODUCTION

1.1 BACKGROUND

Longterm exposure to smoking, sedentary habits, unhealthy diets and possibly stress have been mentioned by the South African Medical Research Council (SAMRC) as risk factors for the development of a group of diseases termed chronic diseases of lifestyle (SAMRC, 2008). The World Health Organisation (WHO) and World Bank Global Burden of Disease Study statistics (1990), as reported by Pauwels et al (2001), showed the estimated prevalence of chronic obstructive pulmonary disease (COPD) in the world as 9.34/1000 in the male population and 7.33/1000 in the female population. The issue of COPD in the elderly was addressed in a study by Buist et al (2007) who concluded that the prevalence of the disease was strongly determined by age and smoking. They suggested that public health authorities would have to take these and other causative factors into consideration when designing preventative measures at a primary and/or secondary level. In a Burden of Obstructive Lung Disease (BOLD) study, looking at the prevalence of COPD in different international populations, 12 study sites were utilised (Buist et al, 2007). A high prevalence of COPD was found in Cape Town, South Africa (Mannino and Buist, 2007). In the South African Demographic and Health Survey (SADHS) (1998) data that emerged from developing countries showed that the incidence of non-communicable diseases (NCD), which are usually linked to a wealthier population, are as high in poorer nations (Schneider et al, 2009). Beaglehole and Yach (2003) pinpointed the fact that, in poorer populations, the rate of globilisation has outstripped the national and international reaction to this phenomenon which resulted in the subjection of struggling health services in such communities to further pressure due to the increase in the occurrence of NCD. Hence Yach (2002) suggested the need for health services to transform in order to meet the chronic care requirements of the poor. His suggestions included increased efforts to decrease risk factors for the development of these diseases with a focus directed towards primary care.

Emphasis was placed on the fact that risk prevention and early intervention in the treatment of chronic diseases would reduce the burden of NCD on healthcare services in developing countries. People in poorer communities tend to be more likely to suffer from chronic bronchitis (Schneider et al, 2009) due to their greater risk of exposure to extraneous smoke from fuels used at home. They tend to be more prone to smoking although the number of cigarettes smoked per capita is less than that seen in the wealthier sector of the population (Schneider et al, 2009). According to the WHO (Assembly document A61/8, 2008) by the year 2020 chronic diseases “are expected to account for 73% of deaths and 60% of the disease burden worldwide.” COPD will rank as the third most common cause of death.
worldwide by 2020 (Heaney, Lindsay and McGarvey, 2007). The WHO “Tobacco Atlas” (2002) showed that males are heavier smokers than females worldwide. Approximately one billion men smoked globally whereas daily female smokers numbered nearly 250 million. A higher percentage of men (50%) in developing countries smoked compared to those in developed nations (35%).

Airflow limitation which is not fully reversible with treatment characterises pulmonary COPD. The progressive nature of the decrease in airflow is linked to abnormal inflammatory processes in the lungs associated with exposure to noxious gas particles. The severity of the disease in some patients may be due to its extra-pulmonary effects. COPD is preventable and treatable (Pauwels et al, 2001; Barnes, 2003; The Global Obstructive Lung Disease (GOLD) guidelines, 2005; Stockley, Mannino and Barnes, 2009). Murărescu, Mitrofan and Mihailovici (2007) stated that COPD is an umbrella term for emphysema, chronic bronchitis and a category of asthma. The signs and symptoms exhibited by patients who have COPD are related to whether the obstruction in the airways is due to emphysema, chronic bronchitis or severe, chronic asthma. Patients with COPD often have a combination of the emphysematous and chronic bronchitis changes in their lungs and their symptoms reflect a mixed pattern of disease. Chronic bronchitis manifests with obstruction of the smaller airways and fibrosis and is characterised by a chronic productive cough which lasts for a period of three months and is present for two years in a row (Murărescu, Mitrofan and Mihailovici, 2007). Emphysema is associated with destruction of lung parenchyma, breakdown of alveolar structures leading to enlarged airspaces, a loss of lung pliability and premature closure of small airways during expiration and forced expiration in particular (WHO, 2005). A chronic, productive cough is found in patients who suffer from emphysema and/or chronic bronchitis. Smoking is described as the main risk factor for the development of COPD (Mannino et al, 2006; Stockley, Mannino and Barnes, 2009). Environmental and occupational exposure to air pollutants factors predispose to an increased risk of developing the disease (Stockley, Mannino and Barnes, 2009). Bronchitis or pneumonia in infancy can predispose to a mean reduction in forced expiratory volume in one second (FEV₁), adjusted for age and height in adult men. These data support a possible causal connection between COPD in later life and lower respiratory tract infections in early childhood (Shaheen, Barker and Holgate, 1995). In non-smokers links have been established between the inhalation of smoke from the combustion of biomass fuels and the development of COPD (Salvi and Barnes, 2009). Passive smoking is noted as a causative factor in the worsening of existing COPD conditions (WHO, 2002). Genetic factors play a role in the propensity for certain individuals to contract COPD. This is mainly due to a deficiency in protease alpha₁-antitrypsin (Janoff, 1983; Mannino and Buist, 2007; Stockley, Mannino and Barnes, 2009). Severity of the disease and the effects that acute
Exacerbations have on the progression of the disease impact on the quality of life (QoL) of patients with COPD (Andenaes et al, 2006; Koblizek et al, 2009; Zhou et al, 2009; Motohashi et al, 2010). Dyspnoea, limited exercise tolerance, fatigue, associated cardiovascular disease, muscle wasting and weakness, osteoporosis, weight loss, depression, and anxiety are all factors which have a bearing on the QoL of patients with COPD (Lisboa et al, 2001; Nazir and Erbland, 2009).

The anti-inflammatory role of exercise is addressed in the literature in relation to land-based exercise (Jankord and Jemiolo, 2004; Kasapis and Thompson, 2005). The link between chronic disease and inflammation is well documented (Zebrack and Anderson, 2002; Pedersen and Bruunsgaard, 2003; Pedersen and Pedersen, 2005). COPD is associated with pulmonary as well as systemic inflammation (Heaney, Lindsay and McGarvey, 2007; Papaioannou et al, 2009). The management of COPD by physiotherapists focuses mainly on land-based treatment methods. Lacasse et al (2009) reported on the effects of pulmonary rehabilitation (PR) on exercise capacity for patients with COPD and the changes in QoL related to health issues in this population. Their Cochrane systematic review demonstrated that pulmonary rehabilitation had a favourable impact on all parameters of QoL measured by the chronic respiratory questionnaire (CRQ) and the St George’s respiratory questionnaire (SGRQ) for patients suffering from COPD. Exercise capacity, as measured through cycle ergometry, showed a positive weighted mean difference of 8.4 watts between the rehabilitation groups and the control groups over the total number of studies appraised. The six-minute walk distance test (6MWD) test, utilised across 16 trials to evaluate functional exercise capacity, demonstrated a weighted mean difference of 48 metres in distance walked between the rehabilitation and the control groups (Lacasse et al, 2009). Langer et al (2009) showed that health-related quality of life (HRQoL) and functional exercise capacity was improved when a clinical practice guideline to treat patients with COPD was implemented.

Immersion in water, in the form of hydrotherapy, has been utilised since the ancient Greek times of Hippocrates (460-377 BC) and the early Roman era (Papavramidou and Christopoulos-Aletra, 2003). Reports of aquatic exercise being utilised in the management of patients with chronic diseases have been recorded in the United States of America between 1840 and 1900. American President Franklin D Roosevelt found that the hydrotherapy he underwent between 1924 and 1927 relieved debilitating muscular spasms experienced as part of his poliomyelitis affliction (Sinclair, 2008). More recently beneficial effects of hydrotherapy on various chronic diseases have been reported. Hydrotherapy was used in the pre-operative management of patients undergoing total hip or knee replacements in order to reduce pain and improve function (Gill, McBurney and Schultz,
In a review, looking at hydrotherapy treatment of a combination of hip and knee osteoarthritis, short term benefits were reported with small to moderate improvements noted in function and QoL and minor improvements in pain levels (Bartels et al, 2007). In a review of information on the treatment of low back pain with hydrotherapy it was concluded that therapeutic aquatic exercise was potentially beneficial in this particular patient population (Waller, Lancbeck and Daly, 2009). Elyan and Khan (2008) reviewed information on the use of water therapy in the treatment of patients with ankylosing spondylitis and reported that the symptoms of the disease, function and perception of health may be improved. Patients with rheumatoid arthritis demonstrated improvements in muscle endurance in both the upper and lower limbs when they performed moderately intensive exercise in warm water (Bilberg, Ahlmen and Mannerkopi, 2005). Hydrotherapy was used to treat pain in fibromyalgia with beneficial outcomes (Evcik et al, 2008). In a randomised controlled trial (RCT) conducted by Chu et al (2004) people with chronic stroke were given a water-based exercise programme. The primary outcome measured was cardiovascular fitness and the researchers concluded that this form of exercise regime, in hemiplegic patients, might be an effective way to promote fitness in that population. They also noted improvements in maximal workload, muscle strength, gait speed and the Borg balance scale score in the intervention group of stroke patients. Noh et al (2008) found improvements in postural balance and knee flexor muscle strength, after an aquatic intervention based on the Ai Chi and Halliwick hydrotherapy methods, in stroke survivors. Water therapy was used to aid recovery, after exercise, in athletes (Banfi and Melegati, 2008) and was also used in an attempt to improve athletic ability (Hamlin, 2007). Dowzer et al (1999) examined the physiological effects of running in deep water and shallow water with a view to seeing whether the information could be extrapolated and applied to recovering athletes in order to maintain their cardiovascular fitness. They concluded that shallow water running (SWR) showed potential as a re-training strategy as SWR elicited higher peak oxygen uptake (VO\textsubscript{2peak}) levels and heart rate (HR) levels when compared to deep water running (DWR) results. They also remarked that it was easier to train in shallow water as there was no need for a flotation device and it was less threatening to patients who had a fear of water.

Beneficial effects of hydrotherapy have also been reported in the elderly as well as in patients who suffer from chronic cardiac conditions. Ide, Belini and Caromano (2005) demonstrated an increase in inspiratory muscle strength with a water-based breathing programme in healthy, aged persons. In patients with chronic heart failure a regime of immersion, gymnastics and swimming led to an increase in cardiac index (Schmid et al, 2007). An aquatic training, detraining and retraining regime in patients with coronary artery disease reported beneficial improvements in VO\textsubscript{2peak}, total body strength and stress test
time during the training and retraining phases (Tokmakidis, Spassis and Volaklis, 2008). In view of the reported beneficial effects of hydrotherapy treatment in the management of patients with various chronic diseases the researcher decided to investigate the role of hydrotherapy as a component of the PR of patients with COPD.

1.2 STATEMENT OF PROBLEM AND JUSTIFICATION FOR RESEARCH

With the increase in the incidence and prevalence of COPD worldwide, as reported in WHO documents and various other surveys, an emphasis has been placed on the search for new treatment options to address effective management strategies for these patients. Pulmonary rehabilitation (PR), in the form of land-based exercises, has been shown to play an important role in improving these patients' QoL by not only addressing their physical needs, but also providing them with interaction with health care professionals as well as other patients with COPD to facilitate the necessary psychological support. Patients with COPD have expressed boredom while doing the same land-based exercises repeatedly and therefore hydrotherapy may offer an interesting alternative exercise option to those patients (Perk, Perk and Bodén, 1996). Much has been written in the literature about cardio-pulmonary changes in healthy individuals, during immersion in water, relative to the effect of temperature and depth of the water on such outcomes and the different exercise modalities employed. It is not known whether much information has been documented on the effects of such interventions on the COPD population. A review of the available literature on the use of hydrotherapy for these patients will be a valuable tool in the investigation, analysis and reportage on the resultant effects of this intervention in the COPD population.

1.3 RESEARCH QUESTION

What is the role of hydrotherapy in the management of patients with COPD?

1.4 SIGNIFICANCE OF RESEARCH

The incidence of COPD is likely to remain high due to the worldwide prevalence of smoking. Compounding the African incidence of this disease is an human immunodeficiency virus/ acquired immunodeficiency syndrome (HIV/AIDS) pandemic that leads to tuberculosis and pneumocystis jiroveci pneumonia and predisposes patients to the development of COPD (Calderon et al, 1996; Probst et al, 2000; Morris et al, 2008) It is feasible to use the proven anti-inflammatory effects of exercise to counteract the inflammatory processes at work in patients suffering from COPD. Land-based exercise in the form of pulmonary rehabilitation programmes (centre-based and home-based) is demonstrated to be beneficial in the COPD population as improvements in QoL, levels of dyspnoea and participation in ADL have been widely reported. However, reported benefits
obtained through hydrotherapy in various patient populations include reduction in pain, increased inspiratory muscle strength and cardiac index as well as several musculoskeletal benefits. The effects of hydrotherapy on the cardiopulmonary and musculoskeletal systems of patients with COPD will be examined in this systematic review to see if hydrotherapy would be a useful additive to existing pulmonary rehabilitation programmes. It is intended that the results could be used by clinicians to set up practice guidelines for the use of aquatic exercise in patients with COPD. It is hoped that the review will assist the researcher to identify areas for future research in the field of rehabilitation of patients with COPD.

1.5 RESEARCH AIM
To systematically review the evidence related to the role of hydrotherapy in the management of patients with COPD.

1.6 RESEARCH OBJECTIVES
- To determine, from the literature, how water temperature and depth of immersion influences cardiopulmonary function in patients with COPD.
- To determine, from the literature, if hydrotherapy increases endurance in patients with COPD.
- To determine, from the literature, if hydrotherapy increases function in patients with COPD.
- To determine, from the literature, if hydrotherapy increases the QoL of patients with COPD.
- To determine, from the literature, whether the anti-inflammatory effects of water-based exercises in patients with COPD have been documented.

1.7 TYPE OF STUDY
A qualitative systematic review of the relevant literature was performed.

Chapter 2 consists of an in-depth discussion of the literature on the development of COPD, its influence on endurance and function in ADL and its effects on QoL. The medical management of patients with COPD is discussed as well as current rehabilitation interventions used in this patient population. Hydrotherapy and its reported benefits in the management of other patient populations are discussed.
CHAPTER 2

2. LITERATURE REVIEW

This review aims to critically evaluate evidence found in the medical and rehabilitation literature about the processes involved in the development of COPD and its impact on the human body. It also evaluates the effects of COPD on endurance, function and QoL as well as benefits and risks reported for hydrotherapy in other patient populations. In addition it aims to provide the evidence to support research into this specific form of exercise in the rehabilitation of patients with COPD.

2.1 FACTORS PREDISPOSING TO THE DEVELOPMENT OF COPD

In a schematic view of lifestyle diseases that share similar risk factors (SAMRC, 2008) emphysema was named as a chronic disease of lifestyle and was linked to smoking. Statistics available showed the prevalence of smoking globally and linked it to the incidence of morbidity and mortality from lifestyle diseases (British Heart Foundation, 2008). In 2001, according to the WHO’s global burden of disease and risk factors project, COPD ranked fifth as a leading cause of death in countries whose citizens were in an upper income bracket and sixth in those countries with citizens in a middle or lower income category. Furthermore COPD accounted for 3-8% of total deaths in the higher income group and 4-9% of total deaths in the middle/lower income group (Mannino and Buist, 2007).

A greater prevalence of COPD was reported in countries where cigarette smoking was common (Pauwels et al, 2001). Bateman and Jithoo (2006) pointed out that the lung was more vulnerable to diseases of lifestyle than probably any other organ in the body. In the South African context they linked this fact to the increase in the prevalence of asthma and COPD caused by smoking. Considering the South African trends, as recorded in the SAHDS of 1998, it was noted, by Schneider et al (2009), that the poor were more exposed to noxious smoke from the fuel used in their homes and were more inclined to smoke than their wealthier counterparts. However, the richer members of society, smoked more cigarettes per capita than the poorer smokers (Schneider et al, 2009). Smoking has been described as the main risk factor for the development of COPD (Mannino et al, 2006; Stockley, Mannino and Barnes, 2009). Biochemical links have been shown to exist between the development of pulmonary emphysema and cigarette smoking with these changes being due to elastolytic activity in the lung connective tissue, acting in an unrestrained fashion and hence causing the breakdown in alveolar structure (Janoff, 1983). Environmental exposure to air pollution was linked to the development of COPD (Stockley, Mannino and Barnes, 2009) as was childhood exposure to lower respiratory tract infections.
(Shaheen, Barker and Holgate, 1995). In non-smokers, particularly in developing countries, links have been established between the inhalation of smoke from the combustion of biomass fuels and the development of COPD (Salvi and Barnes, 2009). Increase in age reportedly leads to an increase in the prevalence, morbidity and mortality of the disease (Mannino and Buist, 2007). Genetic factors also play a role in the propensity for certain individuals to contract COPD. This was mainly due to a deficiency in the protease alpha_{1}-antitrypsin (Janoff, 1983; Mannino and Buist, 2007; Stockley, Mannino and Barnes, 2009).

According to the WHO “Tobacco Atlas” (2002) the prevalence of smoking among males globally was found to be highest in China with 300 million smokers being recorded. This number equalled the entire population of the USA. In the Russian Federation and many of the satellite states which were incorporated in the former Soviet Union there was a reported incidence of smoking of 60% or more among males. On the African continent the highest incidence of smoking among males was in Namibia (50%), Kenya (49%) and Tunisia. The incidence of smoking among men in South Africa was reported to be 42%. Bosnia and Herzegovina and Guinea were shown to have the highest global incidence of smoking within their female populations (WHO, 2002). On the African continent the highest incidence of smoking among females was recorded in Guinea where between 40% and 49% were smokers. In South Africa 11% of the female population were recorded as being smokers. The countries which showed an equal ratio of male smokers to female smokers were New Zealand, Norway and Sweden. Statistics showed that between 1960 and 2000 in the United States of America (USA), United Kingdom (UK) and Japan a gradual decrease in the percentage of smokers in both sexes was recorded (WHO, 2002). Nearly 25% of children who smoke started before the age of ten. In countries such as Ghana, Grenada, India, Jamaica and Poland over 30% of children fell into this category (WHO, 2002). It was reported that the earlier the age at which a person began to smoke the greater the risk would be for that person to contract smoking-related diseases in later life (WHO, 2002).

The prevalence of passive smoking worldwide was such that 40% of children were exposed to it in their home environment as children are at particular risk from adult smoking habits. Pneumonia, bronchitis, coughing and wheezing, asthma attacks and cardiovascular disease can all result from such exposure (WHO, 2002). The reasons that young people start to smoke are linked to tobacco advertising, availability of cigarettes and low prices which made smoking more appealing and accessible. Peer pressure from friends and family often persuaded the child, with a low self image, to begin smoking to raise their social profile. If the parents of children were smokers those children had an increased propensity to smoke (WHO, 2002).
In 1976 a research report by British American Tobacco (WHO 2002) said that women were more highly motivated to smoke than their male counterparts and therefore they were less likely to quit smoking. The liberated, slim, sophisticated and modern woman was what the tobacco industry falsely associated with the image of the female smoker in order to entice young women to begin smoking. In the case of male smokers the lure of smoking was probably due to the portrayal of smokers as being manly, masculine, and wealthy, with increased sexual prowess, health and happiness. The reality was debilitating diseases such as atherosclerosis, cardiac disease, COPD, osteoporosis and anaemia as well as the resultant complications and symptoms of such diseases; stroke, hypertension, dyspnoea, fatigue, decreased exercise capacity, fractures and general decline in the ability to perform ADL and therefore a decreased QoL (WHO,2002). During pregnancy exposure to cigarette smoke was reported to harm the foetus (Ahluwalia, Grummer-Strawn and Scanlon, 1997). In female smokers there was also the risk of babies with a low birth weight and the resultant complications (Dejmek et al, 2002). There was also a proven link between mothers who smoked and babies who suffered from sudden infant death syndrome (SIDS) (Dejmek et al, 2002).

Smoking is one of the most difficult addictions to overcome. According to the WHO “Tobacco Atlas” (2002) the worldwide incidence of those who manage to quit smoking was not very high. The highest percentage of people to quit smoking by 2002 was in Zambia (72%). However, this statistic did not reflect the global trend where the next best percentage for quitting was 42% in the USA. The other recorded values lie in the range between 1% and 40%. Fortunately, in many developed countries (USA, UK, Australia and Canada) smoking among women was on the decline. However, there was a stagnation or increase in smoking among women in southern, eastern and central European countries.

Many different forms of cessation aids to enable the smoker to quit have been marketed over the years. Among these are nicotine-based chewing gums, nicotine-based ‘patches’, cigarette filters which remove a good deal of the noxious gaseous components of the cigarette smoke and more recently a smokeless electric cigarette which delivers a certain quantity of nicotine and fulfills the smoker’s need to have a cigarette to hold and ‘smoke’. In the case of the electronic cigarette the nicotine ‘dosage’ is decreased gradually. Hypnotherapy and motivational methodologies are also applied to enable the smoker to quit. However, the fact still remains that for a smoker to quit successfully they have to use their own willpower to overcome the craving.
2.1.1 The Role of Cellular Mediators in the Development of COPD

Various authors define COPD as a disease characterised by non-reversible progressive airflow limitation with chronic inflammatory changes as a result of exposure to noxious particles and gases (Pauwels et al, 2001; Barnes 2003). Shapiro (1999) investigated the role of the macrophage in COPD. He described the inflammatory cell recruitment in the lung tissue which was initiated by chronic cigarette smoking and resulted in more than five times the number of total cells present during bronchoalveolar lavage (BAL). He reported that 95%-98% of these cells were macrophages and that the elastolytic proteases released by these inflammatory cells destroyed the lung parenchyma, a fact verified by Russell et al (2002). Neutrophils, acting alone or together with alveolar macrophages, produced an increase in protease production (Barnes, 2003). Di Stefano et al (1998) found that there was a correlation between the number of neutrophils in sputum and the degree of severity of the COPD. In patients with emphysema, macrophages were found in the alveolar wall where destruction of tissue was observed (Finkelstein et al, 1995; Mishi et al, 2002). It was hypothesised that the release of inflammatory mediators, such as interleukin-8 (IL-8), TNF-alpha, leukotriene B₄ (LTB₄) and reactive oxygen species, by macrophages stimulated by cigarette smoke, highlighted the link between cigarette smoking and inflammation in COPD (Barnes, 2003). Keatings et al (1996) did a study in which they induced sputum production in 16 non-smoking control subjects, 23 patients with asthma, 12 healthy cigarette smokers and 14 patients with COPD. The researchers concluded that there was evidence to perhaps substantiate the fact that TNF-alpha and IL-8 played a role in the inflammatory processes seen in COPD. Other researchers reported that it was possible that T-lymphocytes played an active role in the pathological changes towards emphysema seen in smokers (Finkelstein et al, 1995; O'Donnell et al, 2005). The combination of cigarette smoke medium (CSM) and TNF-alpha at below threshold levels seemed to increase IL-8 production greatly, a factor not realised when either of these two agents were present alone (Sarir et al, 2009). Citing the work done by Finkelstein et al (1995) and O'Shaughnesssey et al (1997), Barnes (2003) reported that, within the T-lymphocyte cellular category, CD8⁺ T-cells showed the most prolific increase in numbers in the lung parenchyma and central and peripheral airways of patients with COPD. There was a positive correlation found between the extent of alveolar destruction and the limitation of airflow due to obstruction and the number of T-cells.

Saetta et al (1993) showed that, in patients with chronic bronchitis, there was a significant increase in T-lymphocytes and macrophages found in the bronchial mucosa. These researchers noted that, in the case of chronic bronchitis, it was the central airways that were inflamed. MacNee (2005) reported on the findings of a study carried out by Hogg
in which that researcher had stated that the inflammation, in the case of chronic bronchitis, was in the central airway epithelium and also in the mucus-producing glands. Hogg (2004) linked mucus hypersecretion to the inflammatory processes. MacNee (2005), by way of debate on this matter, reported on work done by Vestbo and Lange (2002), and noted that those authors had concluded that early manifestations of mucus hypersecretion and its link to limitation of airflow in COPD was uncertain. Vestbo and Lange (2002) also concluded that, due to the fact that smokers with normal lung function but with increased mucus production do not necessarily develop COPD later in life, mucus hypersecretion appeared to contribute little to the inflammatory processes in the early stages of COPD. MacNee (2005), reporting on a study by Fletcher and Peto (1977), further noted that in the later stages of the disease exacerbations linked to mucus hypersecretion may lead to decreased FEV₁ levels and may be a result of the inflammatory response in the submucosal glands. Hence it appears that MacNee (2005) agreed with the link established by Hogg (2004) between mucus hypersecretion and the inflammatory processes seen in COPD. However, MacNee (2005), after reviewing the work done by Fletcher and Peto (1977), suggested that the inflammatory processes were seen in the later stages of COPD. Saetta et al (1996) showed that, in exacerbations of chronic bronchitis, there was a similar eosinophil expression as in asthma. However, they found an increased amount of interleukin-5 (IL-5) protein in the bronchial mucosa of the asthmatics which was not present in the patients with chronic bronchitis. Inflammation in asthma was mainly eosinophilic and was driven by CD4⁺ lymphocytes (WHO, GOLD global initiative for COPD, 2005; O'Donnell et al, 2005).

Gan et al (2004) undertook a systematic review and meta-analysis to establish the relationship between COPD and systemic inflammation. The authors concluded that C-reactive protein, leukocytes, fibrinogen and TNF-alpha levels were raised in individuals with obstructed airflow. In the case of ex-smokers there was still a degree of low grade systemic inflammation detected. These researchers noted that, in the articles under review, there was mention of a link between systemic inflammation and complications such as osteoporosis (Biskobing, 2002; Agusti et al, 2003), loss of body weight (Di Francia et al, 1994; de Godoy et al, 1996; Eid et al, 2001), pulmonary cachexia (Schols, 2002; Agusti et al, 2003) and cardiovascular disease (Friedman, Klatsky and Siegelaub, 1976; Hole et al, 1996; Schünemann et al, 2000). In a review of studies looking at the possible origins of systemic inflammation in COPD Wouters (2005) cited work done by; Dietrich et al (2002) and Dietrich et al (2003) who reported that, in passive smokers and those with only a few pack years, dysfunction was found in the peripheral vascular endothelium and systemic oxidative stress was recorded; Vernooy et al (2002) looked at the possible role of local pulmonary inflammatory response in the creation of a systemic inflammatory response;
Michel et al (2001) showed that the systemic inflammatory response was independent of that in the pulmonary compartment; it was suggested that hypoxia was a possible reason for the activation of the TNF system and to this end Takabatake et al (2000) found a significant correlation between the prevalence of TNF-alpha in the peripheral circulation and hypoxia in the tissues. One of the key inflammatory indicators found in COPD was C-reactive protein (Vestbo, 2007). He also reported that there was a link between the systemic co-morbidities associated with COPD and systemic inflammation. Cytokines recruited, activated and promoted the ongoing multiplication of inflammatory cells throughout the respiratory tissue (Barnes, 2008).

Therefore in summary, links have been firmly established between smoking and the production of cells which promote inflammatory processes resulting in destruction of lung parenchyma in patients with COPD. When acting alone, or in combination, macrophages and neutrophils promote the release of electrolytic proteases which destroy the tissue of the lung. Macrophages have also been linked to the release of inflammatory mediators TNF-alpha, IL-8, LTβ4 and reactive oxygen species when stimulated by cigarette smoke. Furthermore, CSM in combination with below threshold levels of TNF-alpha seems to increase the levels of IL-8 greatly. T-lymphocytes, and in particular CD8+ T-cells, showed prolific increases in numbers in patients with COPD and contribute to the pathological changes towards emphysema in smokers. Systemic inflammation, found in patients with COPD, was linked to raised C-reactive protein, leukocytes, fibrinogen and TNF-alpha levels. A degree of low grade systemic inflammation was seen in ex-smokers.

2.2 CLASSIFICATION OF SEVERITY OF COPD

The GOLD classification of severity of COPD (GOLD, 2005) states the following: Stage 0 denotes those patients at risk of developing the disease who exhibit a productive cough and presents with normal spirometry; Stage 1 is classified as mild COPD with FEV₁ ≥ 80% predicted; Stage 2 is moderate COPD with 50% ≤ FEV₁ < 80% predicted; Stage 3 is severe COPD with 30% ≤ FEV₁ < 50% predicted and Stage 4 is very severe COPD with FEV₁ < 30% predicted or FEV₁ < 50% predicted with chronic respiratory failure. Bateman et al (2004), in a guideline for the management of COPD in South Africa, linked the stages of severity to dyspnoea and functional impairment. According to these authors Stage 0 is related to normal exercise tolerance; patients with Stage 1 experience limitations with strenuous activity; those with Stage 2 manage limited activities done at a normal rate and those with Stage 3 have a compromised lifestyle with their ADL seriously affected and/or a completely sedentary lifestyle. With regard to performance on the 6MWD test, the distances that patients would manage to walk were as follows: Stage 0 >600 m (normal result); Stages 1 and 2 <600 m-200 m and Stage 3 <200 m. Measures of severity reflect the
progressive nature of the disease and the impact it has on the lifestyle impairment of the patient.

The multi-factorial prognostic index (BODE) which comprises measures of body mass index, airflow obstruction, dyspnoea and exercise capacity was described in a study which compared it to the GOLD classification for the parameters of anxiety and depression in patients with COPD (Funk et al, 2009). The researchers found that the prevalence of anxious symptoms, linked to dyspnoea, increased with the BODE stages of severity of disease but did not increase with the GOLD classification of the stages of COPD. However, with symptoms of depression, explained by both dyspnoea and exercise capacity, the increase shown related to both the GOLD and the BODE scales. The researchers concluded that the BODE index was a better measure than the GOLD classification for explaining depressive and anxious symptoms in patients with COPD (Funk et al, 2009).

2.3 THE INFLUENCE OF COPD ON ENDURANCE AND FUNCTION IN ACTIVITIES OF DAILY LIVING

Skumlien et al (2006) looked at gender differences, in patients with COPD, in relation to ADL. They used the Pulmonary Functional Status and Dyspnoea Questionnaire (PFSDQ) to measure dyspnoea during ADL, general dyspnoea and loss of performance related to function. They reported that the ability to perform ADL was limited by COPD a fact confirmed by Lahaije et al (2010). Skumlien et al (2006) also noted that dyspnoea scores were not influenced by gender differences and that the PFSDQ could be used to measure outcomes accurately in both females and males. They emphasised that the loss of functional performance in females, in relation to household management, had to be considered when designing an effective treatment programme for female patients with COPD.

Garrod et al (2000) stated that the assessment of the level of restriction of ADL in patients with COPD had to report on the impact of the handicap and disability on the ADL. One of the first measures of dyspnoea, developed in association with the Medical Research Council (MRC) Committee on the aetiology of chronic bronchitis, was the Fletcher (MRC) dyspnoea scale (Fletcher, 1960). Dyspnoea and fatigue associated with COPD have led to functional limitation and disability (Pitta et al, 2006). In fact, dyspnoea in patients with COPD was frequently reported in the performance of everyday tasks (Restrick et al, 1993; Bestall et al, 1999). Patients with COPD are far more handicapped in relation to their ability to lead an active life than healthy elderly people (Pitta et al, 2005). The increasingly poorer performance of ADL by patients with COPD was linked to the deterioration in health of
these individuals and the decrease in related exercise capacity (Oga et al, 2005). Pitta et al (2005) looked at a comparison of physical activities in the daily lives of patients with COPD (GOLD classification 2 –4) and healthy individuals in a similar age group. The physical activities assessed included walking, cycling, standing and sitting. They concluded that most patients with COPD spent more time lying down or sitting than walking or standing in daily life. These researchers also noted that the gait of patients with COPD was significantly slower than that of healthy individuals with a lower movement intensity displayed. The authors found that the 6MWD test provided an accurate measurement of the difference in performance between the two groups. It was noted by the researchers that the healthy individuals and patients with COPD who walked more in daily life walked further during the 6MWD test. The mean percentage predicted distance walked by the COPD group was 62 ± 22m and that by the healthy group a significant 98 ± 10m. Sixteen out of 20 patients with COPD, who walked less than 400m, had an average daily walking time of less than 30 minutes. The authors concluded that patients who performed poorly on the 6MWD test were likely to undertake very little physical activity in daily life.

Many methodologies and tools have been developed in order to assess different parameters related to the measurement of the ability of patients with COPD to perform ADL. The American Thoracic Society (ATS) issued guidelines in 2002 for the utilisation of the 6MWD test. It was suggested that the test could be used to assess the effect of a course of pulmonary rehabilitation on exercise endurance when measurements were taken pre and then post treatment. Schönhofer et al (1997) looked at the role of a pedometer as a repeatable measure of evaluating daily activity in patients with COPD. They concluded that such an intervention could be used in conjunction with a questionnaire in order to assess exercise limitation caused by COPD. Van Helvoort et al (2007) conducted a study in which they evaluated the effect of a 6MWD test on patients with COPD with muscle-wasting. Oxidative stress and systemic inflammation were shown to play a part in the systemic immunological response which the authors found to be comparable to that of a maximal cardiopulmonary exercise testing response. They showed a possible interaction between oxidative stress, systemic inflammation and the degree of muscle wasting in patients with COPD and the effect these factors had on ADL. Borel et al (2010) developed the six minute stepper test (6MST) as an alternative to the 6MWD test and suggested it could be used effectively to evaluate exercise tolerance in patients with COPD. It also negated the necessity for a passage of a specific length to have to be used to calculate the data as the stepper took up very little space.

Lacasse et al (2009) highlighted the importance of the demonstration of physiological rationale behind programmes of exercise training in patients with COPD in order to
accurately measure rate of perceived exertion (RPE) and to reproduce these results for comparative purposes. The scale most widely used to assess RPE was the Borg scale (first developed by Borg and Lindblad in 1976). The original scale consisted of 15 grades ranging from 6 to 20 with the configuration being chosen to best match heart rates ranging from 60-200 beats per min^{-1}. However, age, type of exercise, anxiety, environment and other factors influenced the outcomes. The modified Borg scale was constructed as a category scale with ratio properties and ranged between 0 to 10 with 0 being nothing at all and 10 being almost maximal exertion (Borg, 1982).

Dyspnoea was defined by the ATS (1999) as "a subjective experience of breathing discomfort that consisted of qualitatively distinct sensations that varied in intensity." The experience of shortness of breath is a multi-factorial one which resulted from numerous psychological, physiological, environmental and social factors acting together to produce resultant behavioural and physiological reactions (Hill et al, 2004). Borg was also instrumental in the development of a dyspnoea rating scale (Borg, 1982). Another measurement tool, used to measure the intensity of dyspnoea, is the visual analogue scale (VAS). Von Leupoldt et al (2006) compared the VAS scale to the Borg scale in a study looking at the sensory and affective aspects of dyspnoea. Their results demonstrated that the Borg scale was influenced differently by sensory and affective aspects of dyspnoea and hence they concluded that different scales should be considered in the evaluation of the perception and reporting of dyspnoea. Mahler and Horowitz (1994) evaluated these tools in their study and found them to produce valid, responsive and reliable measurements of dyspnoea and efficacy of therapy.

Bestall et al (1999) demonstrated the validity of the MRC dyspnoea scale as a valid means of dividing patients with COPD into categories according to their disability rating and according to the level of severity of their disease. Tangri and Woolf (1973) postulated that it was unlikely that the degree of dyspnoea witnessed in a patient with COPD after a simple ADL like brushing teeth was due to the minor exertion involved. They therefore looked at breathing patterns in patients with COPD and concluded that they were different from those seen in normal control subjects. The main difference seen was that, whereas the patients with COPD had to hyperventilate to normalise blood gas levels post-activity, the healthy group only needed a few deep breaths to achieve the same result. In a study which was prompted by patients with COPD complaining of tiredness and dyspnoea Velloso et al (2003) looked at four ADL, which involved the use of the upper limbs and thoracic muscles, with respect to ventilatory and metabolic outcomes in patients with COPD. The activities were associated with a high energy demand and consisted of erasing a blackboard, replacing lamps, sweeping and lifting pots. The performance of the four activities increased
the oxygen uptake (VO$_2$) demands in the patients and the researchers suggested this was possibly the reason why the patients felt tired during the execution of these easy tasks involving the upper limbs.

Garrod et al (2000) developed the London chest activity of daily living scale (LCADL) in order to validate a standardised measure of ADL in patients with severe COPD. They compared the results obtained with their questionnaire with the activity and impact components of the SGRQ (Jones, Quirk and Baveystock, 1991) and the Nottingham extended activity of daily living questionnaire (NEADL) and found good correlations with both of these. This study highlighted the fact that carers and relatives were necessary to assist patients with severe COPD in their ADL. Disability was marked in the case of self-care items like getting dressed, drying after a bath, putting shoes and socks on and washing hair. Assistance was needed with activities like cleaning a floor or washing curtains. The COPD activity rating scale (CARS) was developed by Morimoto et al (2003). Their results suggested that the scale was a valid tool to use when assessing ADL in patients with COPD. Pitta et al (2006) compared and discussed the use and effectiveness of questionnaires and motion detector devices in the accurate assessment of performance in ADL and functional capabilities in patients with COPD. They found the motion detector type of recording device to be of more value.

Hence, from the above information it can be concluded that COPD has been shown to limit the performance of ADL in patients who are suffering from this disease. The severity of COPD correlates directly with a reduction in exercise capacity and patients with COPD are more likely to follow a sedentary lifestyle. Dyspnoea and fatigue lead to functional limitation and disability. Disability related to self-care matters can be marked and often necessitates the involvement of carers and relatives to assist the patient with such matters. Many measurement tools are available to assess the level of disability related to the performance of functional activities and the limiting factors associated with the ability to perform the ADL. Gender related differences should be considered when drawing up a treatment programme for patients with COPD.

**2.4 THE INFLUENCE OF COPD ON QUALITY OF LIFE**

The term QoL was described by Engström et al (1996) as an indication of a patient with COPD’s mental and physical well-being in relation to the severity of their disease. These researchers compared results from a COPD population and a control group in relation to their performance on the Sickness Impact Profile (SIP), the Hospital Anxiety and Depression scale (HAD) and the Mood Adjective Check List (MACL). They looked at how the results obtained related to pulmonary function, dyspnoea, mental factors and the
indications of the severity of the disease as well as the correlation between the results and smoking habits. The SIP was used to determine the functional status of the patients. They concluded that, in the early stages of the disease patients were able to use coping mechanisms and possibly reserve pulmonary capacity in order to counteract any loss of pulmonary function and hence the generic questionnaires relating to QoL did not ascertain a significant deterioration in this regard. However, in the later stages of the disease, generic questionnaires could be used to discern changes in QoL in patients with different levels of pulmonary dysfunction. Smoking habits were not found to impact on a feeling of well-being. The authors also noted that clinical parameters may not adequately address the specific problems encountered by the individual patient. The SIP is usually regarded as a sickness-related measure however, Engström et al (2001), viewed it as a generic self-assessment questionnaire used to compare important functional aspects between diseases. In this particular application they suggested it was a valuable tool to accurately assess patient-specific problems in order to utilise that information to draw up individualised rehabilitation programmes and assess treatment effect. Curtis, Deyo and Hudson (1994) and Kaplan and Reis (2007) stated that the term QoL was strongly linked to factors other than those dictated by the health of the patients. These included socio-economic factors (financial status, housing, employment and social support) and functional status. They also felt that, because many of the COPD patients were elderly HRQoL was a more fitting terminology when looking at the well-being of this category of patients.

The ATS statement (1995), as reported by Ferrer et al (1997) on the diagnosis and care of patients with COPD stated that Stage 1 COPD minimally affected QoL but stages 2 and 3 were linked to great deterioration in QoL. Ware and Sherbourne (1992) reported that HRQoL in COPD was better delineated by using the SGRQ and the medical outcomes short-form 36-item (SF-36) questionnaire than by the ATS’s recorded stages of the disease. Lisboa et al (2001) linked the severity of impairment in HRQoL, in patients with COPD, to dyspnoea, muscle wasting, the inability to exercise and fatigue. The level of dyspnoea found in patients with COPD was thought to be a useful indication of HRQoL of these patients (Hajiro et al, 1999). Psychological factors such as anxiety and depression were also found to affect the HRQoL of patients with COPD (Lisboa et al, 2001). Ferrer et al (1997) showed that patients with mild COPD had a compromised HRQoL. They also linked the HRQoL to co-morbid disease processes. Verrill et al (2005) looked at how short-term and long-term pulmonary rehabilitation (PR) impacted on QoL. They used the SF-36 questionnaire and the Ferrans and Powers quality of life index-pulmonary version III (QLI) to assess the effects of PR programmes of different duration on the QoL of patients with COPD. They concluded that QoL measures had improved after 12 weeks of PR participation and these improvements were maintained by 24 weeks of PR participation.
Hassanein and Narsavage (2009) found that physical and mental QoL, in patients with COPD, improved clinically after a PR programme of 90 minutes a day for two to three days per week over a maximum duration of three months (24 sessions) and that optimal beneficial outcomes were related to greater patient compliance with the supervised PR programme. Motohashi et al (2010) showed, with the aid of high-resolution computed tomography (HRCT), that it was the severity of emphysema that was associated with HRQoL and generic QoL due to a reduction in diffusion capacity. They suggested that, in clinical practice, HRCT could be useful as a diagnostic tool for use in the elderly patient who was unable to undergo spirometry.

The measurement of the QoL of patients with COPD is subjective. Among the many methods of assessment are: the SGRQ (Jones, Quirk and Baveystock, 1991), the CRQ (Morgan, 1991; Yohannes et al, 1998), the breathing problems questionnaire (BPQ) and the brief assessment of depression cards (BASDEC) for a measure of psychological well-being (Yohannes et al, 1998). The SGRQ was described as a disease-specific health measure by Okubadejo, Jones and Wedzicha (1996). They reported that the relationship between hypoxaemia and the well-being of the patient with COPD was better established through the use of the SGRQ than with the SIP and the quality of well-being (QWB) scales. The SIP was originally developed by Gilson et al (1975) as a health status index and changes from one application to another were meant to provide indications of changes in health in the population being studied. Blinderman et al (2009) used the SIP scale as part of a multi-variate analysis to assess sickness-related impairment across 68 items categorised into social behaviour, somatic autonomy, mobility control, social behaviour and feelings, mobility psychological autonomy and communication.

Kaplan (2005) looked at three published studies of patients with COPD where data were collected and evaluated before and after PR using the QWB scale in order to establish the minimally clinically important difference (MCID). He reported that all three evaluations of PR showed changes on the QWB scale in excess of the predicted MCID value and he concluded that utility-based QoL measures provide data which could be easily interpreted and the methodology was sensitive enough to suggest beneficial effects of such rehabilitation in patients with COPD. In a study done in the USA which looked at HRQoL indexes from the National Health Measurement study the QWB self-administered scale was used as a measure of HRQoL in older US adults along with five other HRQoL measurement tools. The six scales showed similar patterns in age-related HRQoL by gender but the trends relating to HRQoL in the different population norms were similar but not identical. The QWB scale was comparable in its detection of the trends under review (Fryback et al, 2007).
Guyatt et al (1987) were responsible for the development of the CRQ which determined the effect of treatment on QoL in clinical trials. The finalised CRQ examined four aspects of the lives of patients: dyspnoea, fatigue, emotional function and mastery. The latter involved the patient’s ability to deal with the disease and its effects. Morgan (1991) utilised the CRQ in his study in order to measure QoL in patients with chronic lung disease. He reported that he found the test of particular value in showing changes in disability with respect to elderly patients with COPD. In a very recent study Reda et al (2010) looked at the reliability of the chronic respiratory questionnaire self-reported (CRQ-SR) and the clinical COPD questionnaire (CCQ). These researchers concluded that both the CCQ and the CRQ-SR were equally valid and reliable but over the longer term (when the follow-up exceeds 26 weeks) the responsiveness of the CCQ was better in the assessment of QoL factors in patients with mild to moderate COPD.

The deterioration in the QoL of patients with COPD is linked to the systemic inflammatory processes associated with COPD and co-morbid conditions (Gan et al, 2004). As a patient with COPD is likely to suffer from a range of physical and psychological manifestations of the disease Nazir and Erbland (2009) recommended that a multi-disciplinary approach to the management of these patients would be advisable to help them attain a better overall QoL. Fan et al (2002) found that the lower the QoL rating of patients with COPD the higher the likelihood would be for them to need hospitalisation or to die from their condition.

2.5 MEDICAL MANAGEMENT OF PATIENTS WITH COPD

Bateman et al (2004) formed a working group that compiled a guideline for the management of patients with COPD on behalf of the South African Thoracic Society (SATS). The guideline suggested that, before treatment was started, confirmation of diagnosis and the stage of severity of the disease had to be determined by means of spirometry. Based on the severity of the disease, any symptoms of dyspnoea and airflow obstruction are treated. Rehabilitation and education were implicated as important non-pharmacological interventions in the management of these patients. The guideline stated the importance of prevention of exacerbations and prompt treatment if exacerbations did occur. Any co-morbid complications had to be treated and great emphasis was placed on education on cessation of smoking. Similarly, Calverley et al (2007) and Gold (2009) state that treatment for COPD should be focussed on relieving symptoms, decreasing risk factors, preventing exacerbations and reducing mortality.

Short -acting β₂-antagonist (inhalers) or anticholinergic bronchodilators (inhalers), either alone or in combination, are suggested for use in Stage 1 disease (Bateman et al, 2004; Heaney, Lindsay and McGarvey, 2007). Another option mentioned is oral theophylline in a
sustained release form (Gold, 2009). In stage 2 disease regular use of short acting or long-acting bronchodilators alone or in combination with oral theophylline is recommended (Bateman et al, 2004). With increasing severity of disease two or three of these drugs in combination is advocated. Formoterol and salmeterol are the recommended β₂-antagonists and it is suggested that long-acting anticholinergic tiotropium should also be used (Heaney, Lindsay and McGarvey, 2007; Gold, 2009). If FEV₁ is below 50% predicted, oral or inhaled corticosteroid for a trial period is indicated. With three or more exacerbations in a year or if the trial treatment proved effective a maintenance dosage of inhaled corticosteroids or a low-dosage of oral corticosteroid treatment is deemed effective as a maintenance measure (Bateman et al, 2004). In an editorial which looked at pharmacological interventions to reduce inflammation in COPD, Barnes (2007) cited a study by Bourbeau et al (2007) which reported on the anti-inflammatory effects in the airways of patients with FEV₁ ≥ 25% predicted (post-bronchodilator) from salmeterol/fluticasone propionate. They concluded that a combined therapy using a long-acting β₂ agonist together with an inhaled corticosteroid (ICS) was beneficial and had more effect on inflammation than the sole use of inhaled corticosteroids. Calverey et al (2007) looked at the same combination therapy in relation to spirometric values, frequency of exacerbations, health status and survival rates in patients with COPD. They concluded that the level of statistical significance attained, with regard to a reduction in death in the patients with COPD, was not equal to the predetermined level set. However, all other outcomes measured in these patients showed significant benefits. While ICS was shown to potentially reduce the risk of lung cancer in patients (Parimon et al, 2006) the TOwards a Revolution in COPD Health (TORCH) study (Crim et al, 2009) found that ICS alone or in a combination therapy could possibly be linked to the development of pneumonia in patients with COPD. Postma and Calverley (2009) concluded that the use of ICS, either alone or in combination with long-acting β₂-antagonists (LABA) had beneficial outcomes in patients with moderate to severe COPD. In contrast Suissa and Barnes (2009) put forward the case against the use of ICS in the treatment of COPD. They pointed out the RCTs conducted to investigate the validity of the use of ICS had not reported any real value in that form of treatment until the ICS treatment was used in combination with a LABA.

2.6 REHABILITATION INTERVENTIONS FOR COPD

Nlci et al (2006) reported that, according to the American Thoracic Society/ European Respiratory Society (ATS/ERS) statement of 2006, PR is defined as “an evidence-based multidisciplinary and comprehensive intervention for patients with chronic respiratory disease who are symptomatic and often have decreased daily life activities.” Ries (2008) supported this definition and recommended that a tailor-made PR management plan should be drawn up for each patient. A multi-disciplinary approach was proposed in order to
address the physical, social and psychological aspects of the patients’ requirements. The American College of chest physicians (ACCP) and the American association of cardiovascular and pulmonary rehabilitation (AACVPR), recently updated their original guidelines for PR and made new recommendations for outcomes from comprehensive PR programmes (Ries et al 2007). These recommendations, as outlined in a graphic in appendix A, included the use of exercise training of adequate duration and type and which was timed for maximal effect. The authors recommended that exercise programmes be designed for the lower extremity, upper extremity and inspiratory muscles and geared towards improving the patient’s ability to cope with ADL. The use of anabolic drugs was mentioned and treatment of co-morbid conditions was recommended. Education about the disease and supplementation in the form of oxygen, nutrition and non-invasive ventilation was highlighted. Health care costs, utilisation of resources and future research also had to be taken into consideration.

2.6.1 Pulmonary Rehabilitation for Patients with COPD

Only two Cochrane reviews could be sourced on the topic of pulmonary rehabilitation in patients with COPD. Lacasse et al (2009) reported on changes in QoL shown in the studies they included in their review. They reported on results obtained with CRQ and SGRQ measurement tools and showed that PR was favoured overall when compared to normal community care. These authors reported on maximal exercise capacity where measurement was done using cycle ergometry. They showed that the weighted mean difference in performances between the intervention groups and the controls, over 18 reviewed trials, was an increase of 8.4 watts. In the case of functional exercise capacity the authors reported a weighted mean difference in distance walked of 48m between the controls and the intervention group on the 6MWD test. However, the authors did state that the clinical significance of this result was questionable. Overall the authors of this review concluded that PR exhibited beneficial outcomes in the QoL domains and improved the patients’ emotional outlook and mastery of their disease. In a Cochrane review Puhan et al (2009) looked at the effect of PR after exacerbation of COPD but did not state a definitive timing for the PR intervention after the exacerbation. However, they did mention that it would be during the recovery period. They reported that three of the six reviewed studies showed a significant decrease in numbers of patient re-admissions to hospital and a decrease in the need to treat numbers (NNT) in association with rehabilitation. QoL was greatly improved across all four domains of the CRQ and in the activity limited domains and total score measured by the SGRQ. Exercise capacity was increased from 60-215 m in the 6MWD test and shuttle walk tests. A significant decrease in mortality was also noted. The vast improvement in the outcomes measured, post exacerbation, was attributed to the resultant improvement during PR after the deterioration experienced by the patient during
the exacerbation and a possible modification of patient behaviour due to the chance to educate the patient and implement a home exercise programme (Puhan et al, 2009). Puhan et al (2009) point out that the interpretation of the results obtained in their review could possibly be limited by the small size of the trials under consideration.

A comprehensive and multi-disciplinary approach to PR is seen as the cornerstone of effective management of COPD in order to improve QoL and exercise capacity (Spruit and Wouters, 2007). Troosters et al (2010) stated that PR was instrumental in improving QoL, symptoms of exercise tolerance and reducing the chance of an exacerbation in patients with COPD. Exercise training, focussed on muscle dysfunction as a systematic consequence of COPD, improved joint mobility especially when linked to walking. Physical well-being and possible decrease of the perception of dyspnoea experienced by these patients was also noted (Rochester, 2003; Troosters et al, 2010). The use of lower limb exercise was recommended as a mandatory component of a rehabilitation programme (Ries et al, 2007). These researchers also stated that such exercise performed at higher intensity would result in greater physiological benefits than similar exercise performed at a lower intensity. Rochester (2003) concluded that both high and low intensity training was beneficial in improving exercise endurance even in patients with severe COPD. Patients with COPD walked less than their healthy counterparts. Lower limb activity played a large role in ADL. An objective and subjective assessment of activity after rehabilitation showed an improvement (Walker et al, 2008). Man et al (2009) focussed their review on the effects of PR on the quadriceps muscle in patients with COPD. They concluded that clinically PR resulted in an increase in muscle strength and endurance, reduced the fatigability of the muscles and resulted in a small increase in fat-free muscle mass and cross-sectional area (CSA) of the mid-thigh region. Structurally they concluded that a) fibre-type proportions showed no significant changes, b) there was an increase in muscle fibres in the CSA under examination and c) each muscle fibre type had increased access to the capillary network. There were also metabolic advantages reported. In the case of upper limb exercise, Ries et al (2007) recommended unsupported endurance-type training. The two main types of training mentioned by Rochester (2003) were aerobic (endurance-type) training and strength training. Ortega et al (2002) determined the effect of different exercise training strategies in patients with COPD. They found that 6MWD test was improved most in the group undergoing strength training. However, the group undergoing endurance training and the group doing a combination of both types of training showed more improvement in submaximal exercise capacity. In the case of increase in strength, pertaining to muscle groups used in the exercises, the strength group showed similar levels of improvement to that seen in the combined group and this was more than that seen in the strength only group. All groups registered an improvement in dyspnoea ratings as measured by the CRQ.
Langer et al (2009) reported on the use of interval training and/or endurance training in patients with COPD and concluded that comparable results were obtained when the total amount of work performed was the same. Symptoms of dyspnoea were less during the achievement of high work rates when patients with COPD undertook interval training (Vogiatzis, Nanas and Roussos, 2002; Vogiatzis et al, 2005). Many interventions already described were for group therapy. Serres et al (1997) looked at individualized exercise training in patients with COPD and concluded that, when these patients exercised at the gas exchange threshold (GET) level, peripheral muscle performance was rapidly increased.

Ries et al (2007), reporting on joint ACCP/AACVPR guidelines, stated that PR programmes of a greater duration than 12 weeks showed better sustainable outcomes and the benefits achieved during 6-12 weeks of PR decreased gradually over 12-18 months. HRQoL remained above pre-intervention levels at 12-18 months. Liddell and Webber (2010) looked at a once/week supervised rehabilitation programme versus a twice/week programme and noted similar improvements in exercise tolerance resulting from both. However, there was less improvement in HRQoL in the group that was supervised once a week.

2.6.2 Inspiratory Muscle Training as a Component of Pulmonary Rehabilitation

Garrod and Lasserson (2007) reported that the role of physiotherapists in the treatment of patients with chronic lung disease included the use of a variety of techniques to improve ventilation in order to achieve successful treatment outcomes. Hamilton et al (1995) stated that inspiratory muscle weakness in COPD was known to contribute to dyspnoea. The work done by Gosselink, Troosters and Decramer (1996) found that inspiratory muscle weakness was linked to exercise limitation. COPD was linked to a functional weakness in the inspiratory muscles caused by hyperinflation and the effects of exacerbations of the disease (Larson, Covey and Corbridge, 2002). They also found that the decrease in inspiratory muscle strength was associated with fatigue and dyspnoea. In a systematic review of 18 studies O’Brien et al (2008) compared reported results in relation to inspiratory muscle training (IMT) versus exercise and combined IMT and exercise versus exercise alone. No significant difference in effect was noted in those cases where IMT was compared with exercise. However, the data suggested that a combination of IMT and other exercise interventions over a period of eight weeks at a rate of three times per week could be instrumental in increasing exercise tolerance and inspiratory muscle strength in patients with COPD. In a study, in which patients with COPD, underwent a 12 week general exercise reconditioning programme (GER) followed by a course of IMT or sham IMT with a
GER for six months Magadle et al (2007) found the GER and IMT programme beneficial to outcomes relating to 6MWD test, dyspnoea ratings and SGRQ score.

Specific expiratory muscle training (EMT) was performed by patients with COPD daily, six times a week for 30 minutes per session and over a period of three months. An inspiratory muscle trainer was used with the patients breathing through the expiratory port at a resistance equal to 15% of their maximum expiratory pressure (PE_{max}) for one week. The target of 60% of the baseline PE_{max} was achieved after one month by means of increments of five to ten percent per session. Each week after that further adjustment was made to the new PE_{max} achieved. This intervention resulted in a significant improvement in exercise performance (Weiner et al, 2003a). The same researchers (Weiner et al, 2003b) in a later study in the same year, looked at a comparison of EMT, IMT and a combination of the two treatment methodologies. They found that each individual muscular intervention resulted in an increase in strengthening and endurance in that particular muscle group. All interventions led to improvement in 6MWD test results but IMT and IMT combined with EMT showed greater improvement than EMT alone (Weiner et al, 2003b). Weiner and McConnell (2005) reviewed findings on IMT, EMT and a combination of these modalities in patients with COPD and concluded that strength and endurance can be improved when specific training is given to inspiratory or expiratory muscles; IMT results in an improvement in exercise tolerance, decreases dyspnoea and offers a better QoL; EMT leads to an increase in exercise performance; a combination of IMT and EMT probably results in no additional benefit. Beckerman et al (2005) found that a course of specific IMT over a year benefitted patients with COPD by decreasing dyspnoea ratings and improving QoL. It was also shown to be cost effective as it decreased the use of health services and hospitalization.

2.6.3 Home-Based Versus Hospital-Based Rehabilitation

No Cochrane review was sourced relating to home-based versus hospital based rehabilitation but there were a number of trials found and some will be reported. Larson et al (1999) reported that the effects of IMT and exercise training were previously only demonstrated in the inpatient and outpatient scenarios and so they examined the effects of cycle ergometer training (CET) and IMT in a home-based programme. Patients with COPD were randomised to one of four groups IMT, CET, IMT and CET or health education and were trained for a period of four months. A reduction in exercise related dyspnoea and leg fatigue was recorded with CET. However, this did not translate into a decrease in dyspnoea in ADL. Inspiratory muscle training increased inspiratory muscle strength and endurance. The combination of CET and IMT did not amplify any of the effects noted in the singular interventions. In a clinical trial Battaglia et al (2006) examined the value of using a flow-
volumeric inspiratory exerciser, named Respivol, in IMT undertaken by patients with COPD in the home environment. The patients used the inspiratory exerciser for a total period of 45 minutes per day consisting of three 15 minute sessions and the trial lasted for a period of six months. Patients were assessed before the trial, at the three month stage and at the end of the six month period. The use of the exerciser seemed to have a beneficial effect on the 6MWD test, dyspnoea and QoL ratings after six months of treatment. Ghanem et al (2010), in a randomised clinical trial, evaluated the use of a home-based PR programme undertaken over two months with outpatient supervision every two weeks. The programme consisted of respiratory muscle training (diaphragmatic breathing and pursed lip breathing), Endurance training (walking and cycling) and strength training and isolated muscle training. Significant improvement was seen in the 6MWD test, HRQoL, Arabic translated CRQ (CRQ-SAS) and the SF-36 scores. These authors recommended that questions that needed to be addressed, in relation to PR, were the long-term effects of early PR and the best location, format and length for a PRP.

In summary, it has been shown that in the home –based setting patients with COPD benefit from CET, IMT or an intensive PR programme. It appears that an effective home-based PR regimen can be implemented in order to address some of the problems experienced in the COPD population. However, there appears to be a need to relate such benefits to a better management of ADL in the lives of patients with COPD.

2.7 HYDROTHERAPY
2.7.1 Methods of Hydrotherapy Application
Hydrotherapy is an umbrella term for different kinds of applications of therapy using water in its many forms. There are various methods of hydrotherapy application which include Ai Chi, Halliwick, Watsu and Bad Ragaz. In Ai Chi exercise is performed in 30°C to 35°C (thermoneutral) water up to shoulder level and utilises slow wide movements of the upper extremities, lower extremities and the torso in combination with deep breathing which is reported to benefit the internal organs particularly the heart, liver and lungs (Sziróki, 1998; Kelsey, 2006). The Halliwick concept was first described by James McMillan in England in 1949 and consists of a ten-point motor learning programme which includes acclimatisation to water and rhythmic and rotational movements in water to improve balance. The final stage incorporates swimming with the type of stroke dictated by the patient's abilities and/or disabilities (Sziróki, 1998). Halliwick is usually performed as group therapy but each patient has an individual instructor who assesses and treats the patient accordingly. Breathing techniques are combined with the movements done in the water. The Watsu hydrotherapy method combines the stretches of Zen Shiatsu with the support of the water enabling the spine to move freely with no weight being placed on it and therefore facilitating spinal
muscle relaxation. This relaxation is reported to result in increased circulatory benefits with
the removal of metabolites from the muscles resulting in a reduction of pain and fatigue
(Kelsey, 2006). The Bad Ragaz ring method uses a combination of relaxation, stabilisation
and progressive resistance exercises. The rings serve as flotation and stability aids. In the
application the therapist serves as the fulcrum and is only immersed to the level of vertebra
T9. The patient is supine and the treatment time is gradually progressed to a maximum of
30 minutes. The exercise difficulty may be altered by the control of parameters such as the
length of the lever arm, the range of movement, direction and size of the movement and the
equipment used in the exercise session (Sziráki, 1998).

2.7.2 Effect of Hydrostatic Pressure on the Cardiovascular and Respiratory Systems in
Healthy Individuals

Skinner and Thomson (1983) reported that, when a human body is immersed in a fluid, a
resultant pressure is exerted on every part of the surface area of that body. This pressure,
termed hydrostatic pressure, is exerted by the molecules in the fluid. Hydrostatic pressure
is reported to affect pulmonary pressure balance (Hall, Bisson and O’Hare, 1990; Anstey
and Roskell, 2000). The hydrostatic pressure has both a direct and indirect effect on the
respiratory system. In a review Anstey and Roskell (2000) cited original work done by
Agostoni et al (1966) where they reported that, during the process of chest expansion, the
inspiratory muscles are required to work against this hydrostatic pressure as it causes a
restriction in chest expansion. Anstey and Roskell (2000) also cited early work done by
Jarrett (1965) who found that hydrostatic pressure alters the ratio between the intra-
pulmonary and extra-pulmonary pressure of the thorax in an immersed individual. For the
chest to maintain a relaxed state and a normal resting volume these two pressures have to
be equal. At normal resting volume, the intrapulmonary pressure in the relaxed, vertically
immersed chest is equal to +19.0 cm H$_2$O relative to the pressure at the sternum notch.
Jarrett (1965) also noted that the intrapulmonary pressure is 10.4 cm H$_2$O, relative to the
pressure at the sternum when the patient is immersed in the supine position; and that, in
this same position in air the measurement is 3.4 cm H$_2$O. Hall et al (1990) found that,
during immersion to the supra-sternal notch in the seated position, the intrapulmonary
pressure was +20.0 cm H$_2$O in relation to the hydrostatic pressure. Clinically this factor is
taken into consideration by hydrotherapists when positioning subjects in the water. Different
depths of immersion are used in order to minimise this imbalance. As these facts pertained
to responses observed in healthy individuals, in the case of patients with lung pathology or
rib deformity, treatment modifications were recommended by some authors (Anstey and
Roskell, 2000). In this regard an increase in water temperature (up to 40°C) can lead to an
increase in vital capacity (VC). This could possibly be linked to an increased synergy in
contraction of the respiratory muscles. A change in abdominal wall stiffness, due to the
increased temperature of the water, could lead to an increased abdominal contribution to
respiration (Choukroun, Kays and Varène, 1989). It seems feasible that if weak inspiratory
muscles are unable to contract effectively to facilitate expansion of the chest wall against
the hydrostatic pressure, selective lung volumes would decrease when the patient is
immersed in water (Hall, Bisson and O’Hare, 1990). These authors suggested that the
pathology of a condition could affect the way the patient reacts to extra-pulmonary
pressure. Greenleaf (1984), in a brief review, examined the effect of water immersion on
patients subjected to prolonged bed rest in the horizontal position. He stated that
compression of the abdomen by hydrostatic pressure could act indirectly upon the
diaphragm forcing the diaphragm cephalad resulting in decreased space for thoracic
structures. Agostoni et al (1966), using oesophageal balloons, concluded that there was a
reduction in functional residual capacity (FRC) of 50% due to cephalad movement of the
diaphragm. Agostoni did not take into account whether there were any pressures exerted
from distended mediastinal structures in his calculations and the review by Anstey and
Roskell (2000) reported that his findings remained unchallenged as more recent studies
reviewed had not considered such pressure changes either. Sayliss (1995) and Anstey and
Roskell (2000) suggested that water buoyancy acted like a support for the abdominal
contents resulting in a decreased limitation of respiratory parameters during immersion.

A marked increase in central venous pressure (CVP) was reportedly seen in healthy males
during immersion in thermo-neutral water due to an increase in preload measure of stretch
in the ventricular muscle just before it contracts (Christie et al, 1990; Gabrielsen, Johansen
and Norsk, 1993; Park, Choi and Park, 1999). During upright immersion in water to neck
level it was shown that a movement of venous blood occurred from the periphery of the
body towards the head causing an increase in central blood volume (CBV) (Park, Choi and
Park, 1999). This vascular redistribution into the intra-thoracic vascular bed, of
approximately 700 ml of blood, brought about an increase in pleural pressure in vulnerable
areas and an increase in the weight of the lung. These factors contributed to a reduction in
FRC (Lungren and Fahri, 1989). This increase in the thoracic vascular content also
decreased the elastic recoil of the lungs at low lung volumes (Prefaut, Lupi–H and
Anthonisen, 1976). The combination of the afore-mentioned changes favoured airways
closure with a resultant increase in closing volume (CV) being reported (Bondi, Young et al,
1976; Prefaut, Lupi–H and Anthonisen, 1976). The redistribution of blood volume centrally
resulted in a lowering of HR, in healthy men, compared to that measured in air (Avellini,
head-out graded water immersion to simulate the neural and humoral controlling factors of
cardiovascular functions in healthy subjects under conditions of weightlessness. They
found that cardiac output (CO) was increased with immersion even though the HR decreased. Similar results were reported previously by Hong et al (1969), Arborelius et al (1972) and Begin et al (1976) and Farhi and Linnarsson (1977). Further cardiopulmonary changes observed with head-out immersion were an increase in right atrial pressure when the level of the water was raised from the diaphragm to the neck (Echt, Lange and Gauer, 1974; Risch et al, 1978) accompanied by a resultant decrease in VC due to the increase in the pulmonary circulation (Risch et al, 1978).

2.7.3 The Effect of Different Levels of Immersion on the Cardiovascular and Respiratory Systems in Healthy Individuals

With immersion to the level of the xiphoid process VC, measured in litres, was found to decrease by 7.9% (Greene et al, 1974). This parameter reportedly decreases by 9.9% when immersion is to the level of the larynx (Bondi et al, 1976). During immersion to the neck it reportedly decreases by 8% (Farhi and Linnarsson, 1977; Risch et al, 1978) or by 9% (Agostoni et al, 1966). Greene et al (1974) report that with immersion to the xiphoid process there is a decrease in FRC, measured in litres, of 10.4%. Farhi and Linnarsson (1977) detailed this decrease as being 400ml with immersion to the xiphoid level and a further 400 ml decrease with immersion to the neck. Central venous pressure increased as depth of water increased (Gabrielson, Johansen and Norsk, 1993). Risch et al (1978) recorded a 15% decrease in HR during immersion from xiphoid to neck. However, HR was found to decrease with increasing depth of immersion up to the xiphoid level but then it remained static up to neck level (Gabrielson, Johansen and Norsk, 1993). Central venous pressure increased with a rise in water level (Risch et al, 1978; Gabrielson, Johansen and Norsk, 1993). The former researchers also found the greatest increase occurred when water level was raised from the xiphoid process level to C7. Pulmonary arterial pressure, CO and stroke volume (SV) all increased with increased depth of immersion (Lollgen et al, 1981; Mano et al, 1991). The decrease by 9.4% in residual volume (RV) and by 41.3 % in CV shown by Bondi et al (1976) had an effect on the tidal volume (VT) in 90% of the individuals taking part in their study.

2.7.4 The Effect of Different Water Temperatures on the Cardiovascular and Respiratory Systems in Healthy Individuals

In the cold water study undertaken by Mekjavic and Bligh (1987) an immediate increase in ventilation was noted while sudden cold water stimulus to the skin was referred to as a “gasp” reflex by these authors. McArdle et al (1976) noted a higher VO2 level in subjects immersed in 18°C and 25°C water compared to results in air or thermoneutral water. SV
levels increased significantly in a similar manner to \( \text{VO}_2 \) results in the cooler water temperatures. Šrámek et al (2000) immersed subjects in water at 14\(^\circ\)C, 20\(^\circ\)C and 32\(^\circ\)C in a random sequence with one week between each exposure. They reported that the intervention in 14\(^\circ\)C water led to an increase in HR, SBP and DBP when compared to results obtained with a similar immersion in thermoneutral water. The tidal volume abdominal component (\( V_{T\text{ABD}} \)) was decreased while the tidal volume ribcage component (\( V_{T\text{RC}} \)) was increased during immersion at 25\(^\circ\)C (Choukroun, Kays and Varène, 1989). Choukroun, Kays and Varène (1989), after an intervention where subjects were immersed in thermoneutral water, confirmed the result previously reported by Craig and Ware (1967) where those authors found that VC was decreased significantly when subjects were immersed in water of either 27\(^\circ\)C or 35.5\(^\circ\)C. In their study where an intervention, in which water with a temperature of 40\(^\circ\)C was used, Choukroun, Kays and Varène (1989) found a resultant increase in VC at this temperature. These researchers also show that inspiratory reserve volume (IRV) increases linearly with increments in water temperature while FRC is decreased in thermoneutral water. Increased water temperature has no effect on this parameter. Immersion at 39\(^\circ\)C resulted in an increase in \( \text{VO}_2 \) and CO (Boone, Westendorf and Ayres, 1999). These researchers noted an increase in HR but a decrease in SBP and DBP at this temperature after 15 minutes of immersion.

Thermoneutral water was described as water with a temperature between 32\(^\circ\)C and 35\(^\circ\)C (Choukroun, Kays and Varène, 1989; Nakanishi, Kimura and Yokoo, 1999). However, in an early study Craig and Dvorak (1968) reported that, during dynamic exercise, the temperature of water required for thermoneutrality was lower than that used for static immersion experiments. Water temperatures between 29\(^\circ\)C and 33\(^\circ\)C were therefore recommended for aquatic exercise programmes (McArdle et al, 1976; Choukroun, Kays and Varène, 1989; Choukroun, Kays and Varène, 1990). However, on the basis of the results of their study Israel et al (1989) reported that, during exercise at approximately 60\% of maximum oxygen consumption (\( \text{VO}_2\text{max} \)), immersion in 25\(^\circ\)C and 21\(^\circ\)C water prevented a rise in core temperature. However, the same exercise regimen in 29\(^\circ\)C water led to an increase in core temperature. Hence these researchers concluded that, individuals of similar body composition and fitness level, performing cardiovascular exercise, in 21\(^\circ\)C or 25\(^\circ\)C water temperatures would benefit from the cooling effects resulting from immersion at these temperatures. These researchers also reported that very low intensity hydrotherapy could be performed in 29\(^\circ\)C water as they had found that static immersion at that temperature had resulted in a fall in core temperature in the subjects in their study. Israel et al (1989) concluded further that, the individuals in their study, had a perception of the temperature changes but not an accurate estimation of thermal balance and they therefore recommended that it was important to monitor core temperature during aquatic exercise.
Variations in water temperature change the amount of thermal exchange through the skin (Choukroun, Kays and Varène, 1989). Wilmore and Costill (1994) pointed out that the thermal conductivity of water was approximately 26 times that of air and rapid loss of body heat occurred by conduction and convection during immersion. The passage of water over the body during aquatic exercise further accelerates the heat loss (Datta and Tipton, 2006). Choukroun, Kays and Varène (1990) noted that the temperature of the water had no effect on diaphragmatic contraction but contraction of the more superficial rectus abdominis was subject to the effects of immersion at 25°C and 40°C.

2.7.5 Submaximal and Maximal Workloads in Water for Healthy Individuals

Barbosa, Garrido and Bragada (2007), in a study comparing physiological adaptations to aquatic exercise performed at different levels of immersion, found a higher RPE when the participants exercised with the water up to hip level as opposed to being immersed up to breast level. They concluded that RPE was higher in water up to hip level due to the drag forces associated with the lower limbs being greater than those associated with the trunk and upper limbs. It was reported by Barbosa, Garrido and Bragada (2007) and Barbosa et al (2009) that, with increasing body immersion, there was a decrease in \( V_O{2}_{peak} \) and energy expenditure. The reasoning behind these findings was that: increased depth resulted in an increase in hydrostatic pressure and reduction in cardiopulmonary workout; the action of anti-gravitational muscles in the lower limbs was negated by the buoyancy of the water; and dissipation of heat was greater when the subject was immersed in the deeper water (Barbosa, Garrido and Bragada, 2007; Barbosa et al, 2009).

There are many different types of exercises described in the literature for head-out aerobic exercise programmes. Sanders (2000) categorised these as (a) running, (b) jumping, (c) rocking, (d) scissors (e) kicking and (f) walking. Barbosa et al (2009) stated that, because these basic exercise modalities could be adapted, combined or performed in isolation to many different degrees, it was hard to pin point the actual physiological outcomes to be looked at the different outcomes arising out of a comparison of walking or running in water. In an early study Arborelius et al (1972) found that at rest in thermoneutral water or on land or performing moderate exercise in thermoneutral water or on land HR values were not significantly different. Why and Schoene (1987) conducted a trial which compared HR levels of healthy females performing water walking in waist-deep water and participating in land–based treadmill walking. They reported that water walking elicited significantly higher HR results than land-based treadmill walking. Hall et al (1998) reported that there was an increase in \( V_O{2} \) and HR in healthy female subjects, with an increase in land-based treadmill or water-based treadmill speed. They also recorded that the HR readings were significantly higher when the participants exercised in 36°C water as opposed to 28°C water. Walking in
thermoneutral water while immersed to the level of the xiphoid process resulted in an increase in HR (Shono et al, 2000). Pohl and McNaughton (2003) noted that thigh-deep water walking raised the HR of their student participants significantly more than waist-deep water walking but that running at any depth elevated the HR of the participants the most. In regard to VO$_2$ per stride it appeared that, due to the interaction between water resistance and buoyancy, no difference was seen in the energy expenditure demands for walking or running at any water depth. Oxygen consumption was raised in both walking and running scenarios, at both water levels, compared to land-based treadmill values. The respiratory exchange ratio (RER) was raised during running on land or in water. Pohl and McNaughton (2003) concluded, from the results obtained for HR, VO$_2$ per stride VO$_2$ and RER, that running in waist-deep and thigh-deep water seemed to induce a physiological response that could stimulate a training response with the added value of less impact forces being applied on the lower limbs during such training.

Intense DWR was shown to be effective as an endurance training technique in competitive runners and, due to decreased impact forces, running in deeper water had better rehabilitation outcomes in the initial stages of rehabilitation (Ritchie and Hopkins, 1991). A progression to running in the shallower water to increase the intensity of the rehabilitation programme was recommended as VO$_2$ levels were higher when trained runners ran in shallow water as opposed to running in deep water (Ritchie and Hopkins, 1991; Town and Bradley, 1991; Pohl and McNaughton, 2003). Nakanishi, Kimura and Yokoo (1999) reported that DWR resulted in a decrease in HR. In their review pertaining to physiological changes associated with DWR, Reilly, Dowzer and Cable (2003) reported that SV and CO increased during water immersion and suggested that this was due to the increased arterial blood volume activating the Bainbridge reflex. The resultant tachycardia effect counteracted the expected reflex bradycardia effect (Reilly, Dowzer and Cable, 2003). While DWR training had a sustained aerobic value in trained endurance athletes Reilly, Dowzer and Cable (2003) found that sedentary individuals benefitted more with regard to the improvement in VO$_2$max.

The HR of elderly participants during water walking was akin to that found when the subjects walked on land in contrast to the usual decrease in HR found with immersion of young, healthy individuals (Takeshima et al, 1997). These researchers validated the hypothesis that, when the elderly are immersed in water, the resultant vascular engorgement of the thoracic area leads to a different HR-VO$_2$ relationship than that seen in younger subjects. The authors felt that this fact had to be taken into account when prescribing a water-based exercise regime for the elderly. Takeshima et al (2002) took healthy, elderly women and allocated 50% of the participants to a training group and 50%
to a control group. The water temperature was 30°C and the water level was approximately up to the xiphoid process. The training group underwent a 12 week water exercise programme. The well-rounded exercise programme resulted in a significant 12% increase in VO2peak and VO2 at a 20% lactate threshold (LT). This RCT would have a 1b rating on the hierarchy of evidence ranking system (Lloyd-Smith, 1997) (Table 4.2) and therefore the resultant 12% increase in VO2peak was a significant outcome. As there were no significant changes in the cardiovascular variables in the control group, the researchers concluded that such a water-based exercise programme was beneficial to the cardiovascular well-being of elderly women. Broman et al (2006) looked at the cardiovascular effects resulting from a study on older women where they were required to do land-based treadmill running (TMR) and also DWR in 27°C water. The exercise protocol on the treadmill was paralleled by a programme of DWR where the participants increased the stride frequency to increase the work load in the water. The resultant HR, VO2 level and the ventilation rate was lower in the water at maximal work. However, at submaximal workload, the HR was higher in the aquatic environment. Nikolai et al (2009) investigated whether water aerobics, for a period of 50 minutes in chest-deep water, would give the 14 non-smoking male and female participants an alternative to a land-based exercise programme. The results led the researchers to confirm that the aquatic exercise was adequate to fulfil the required American college of sports medicine (ACSM) criteria for exercise in middle aged or older adults.

2.7.6 The Role of Hydrotherapy in the Management of Chronic Diseases

The use of water rehabilitation in musculoskeletal conditions was covered briefly earlier in this dissertation in relation to treatment outcomes for hip and knee replacements. In a pre-operative population of hip or knee replacement patients Gill, McBurney and Shultz (2009) looked at pain and self-reported function as primary outcomes and timed walk and chair stand as secondary outcomes. The two groups, in the trial, performed a six week programme either land-based or water-based. The regimen included twice-weekly exercise classes. There were no differences reported between the two groups in relation to the primary and secondary outcomes during both interventions. However, immediately after a class there was a more marked decrease in pain associated with the aquatic therapy. Bartels et al (2007), in a Cochrane review, looked at studies that investigated the treatment of hip and knee osteoarthritis with aquatic therapy. No specific time frames were give for the duration of the interventions. The reviewers stated that exercises were undertaken for different lengths of time and for a differing number of sessions per week. They did report that the effects of the different interventions were mainly measured after three months. They concluded that, after an aquatic intervention, there was a small to moderate effect on QoL and function and pain was slightly reduced. They felt that this form of therapy would be
of particular value in the earlier part of an exercise programme for such patients. Elyan and Khan (2008) reported on a small study which demonstrated that hydrotherapy combined with home exercise as opposed to exercise alone had a significant short-term beneficial effect on pain, neck mobility and stiffness in patients with ankylosing spondilitis. Waller, Lambeck and Daly (2009) in a review looked at whether there were beneficial outcomes from an aquatic therapy intervention in the management of patients with low back pain (LBP) or pregnancy-related back pain. They found that beneficial outcomes were reported in most studies and no study in their review reported negative effects of hydrotherapy on the condition of the patient. An aquatic intervention for patients with rheumatoid arthritis was undertaken by Bilberg, Ahlmen and Mannerkopi (2005). The primary outcome measured was aerobic capacity and this was found to be unchanged by the intervention. However, the secondary outcome measured was muscle endurance and hydrotherapy had a beneficial effect on this parameter. Evcik et al (2008) compared an aquatic therapy intervention to a home-based intervention in patients with fibromyalgia. The researchers looked at the effect of the intervention on functional capacity, depression, number of tender points and pain. Both interventions had beneficial outcomes for all four parameters but aquatic therapy resulted in pain relief over a time.

Stroke had also been successfully treated with Bad Ragaz therapy which resulted in a reduction of muscular tone, increase of range of movement (ROM), relaxation, strengthening and muscle re-education (Sziraki, 1998). Chu et al (2004) conducted a RCT to assess water-based exercise for cardiovascular fitness in people with chronic stroke using Ai Chi and Halliwick methods. The experimental group participated in chest deep water exercises over a period of an hour, three times a week. A resultant 22% cardiovascular improvement in the experimental group was reported (Chu et al, 2004). Another trial showed that postural balance and knee flexor strength were improved after such an intervention, over a period of eight weeks, in stroke survivors (Noh et al, 2008).

Tanaka (2009) stated that there was evidence that swimming reduced the chances of regular swimmers developing some of the risk factors for cardiovascular disease. In a RCT of older patients with chronic heart failure the participants were randomised into either eight weeks of a hydrotherapy exercise programme or into a control group. The patients tolerated the intervention in the warm water well with no adverse effects being reported. There was an improvement in maximal exercise capacity, muscle function in small muscle groups and in isometric endurance in knee extension over the results seen in the control group (Cider et al, 2003). Cardiac output was reportedly increased adequately during water immersion, gymnastics and swimming in patients with heart failure. The symptom – limited exercise
stress test, carried out in thermoneutral water, was well tolerated by the patients (Schmid et al, 2007).

2.7.7 The Pathophysiological Effect of Cold Water Therapy
Goedsche et al (2007) addressed the pathophysiological effect of a cold water intervention. The study was not a dynamic exercise intervention and involved the application of cold water stimulations to the upper body. The researchers recorded increases in TNF-alpha, a factor associated with an increased inflammatory response in the body. In frail, elderly humans, resistance exercise had decreased skeletal muscle TNF-alpha according to a study done by Griewe et al (2001). Goedsche et al (2007) cited work done by Rabinovich et al (2003) who had noted increased TNF-alpha plasma levels in COPD patients when they participated in moderate–intensity exercise. Evidence that chronic hypoxia in patients with COPD was related to the activation of the TNF-alpha system was offered by Takabatake et al (2000). These studies possibly highlight the fact that exercise intervention in patients with COPD was found to increase TNF-alpha levels. Elderly individuals who do not have COPD exhibit the opposite reaction as their TNF-alpha levels decrease with exercise (Griewe et al, 2001). The fact that there was an increase in IFN-γ, a cytokine that is responsible for the effective defence against bacteria, and an increase in natural killer (NK)-cells was suggested to play a part against infection by Goedsche et al (2007). They suggested that their intervention was responsible for bringing about an immunological benefit to the patients in their study. Due to the fact that the patients experienced less infections during the follow-up phase and after the study period ended the argument that was put forward by Goedsche et al (2007) that the increased production of the NK-cells, thought to be responsible for this phenomenon, was a slow reactive process.

In summary, no evidence was found in this literature review that a systematic review had been conducted on the topic of hydrotherapy in the management of patients with COPD. Available statistics point to an increase in the incidence of chronic disease (including COPD) worldwide. Aquatic exercise was shown to result in beneficial cardiopulmonary outcomes when used in healthy individuals and in the treatment of inflammatory conditions related to chronic musculoskeletal diseases. It is therefore relevant to investigate its role in the management of patients with COPD as reported by other authors. The next chapter of this dissertation will describe the methodology that was followed in order to conduct the systematic review.
CHAPTER 3

3. METHODOLOGY
The methodology discussed in this chapter is based on the findings of the literature review discussed in chapter 2. The study design, sample population, hypotheses tested, search strategy and data collation used are discussed in detail.

3.1 STUDY DESIGN
A qualitative systematic review was performed.

3.2 RESEARCH HYPOTHESES
- It is hypothesised that there is enough pertinent, up-to-date and clinically significant information available in the literature to produce a systematic review on the topic of hydrotherapy in the management of COPD.
- It is hypothesised that the anti-inflammatory effects of exercise in land-based physical training apply to exercise in water.
- It is hypothesised that hydrotherapy can be used with beneficial outcomes in the COPD population.

3.3 SEARCH STRATEGY
A search was made for relevant articles in the PUBMED, SCOPUS, MEDLINE, CINAHL and SPORTSDiscus databases and the Cochrane review database. Google scholar was perused in order to find any grey literature pertaining to the population under review. The review of the literature was from 1996 until 2009. This protracted period of 13 years was needed to ensure that all leading articles on the subject under review were included in the review. There were too few articles available from 2000 onwards to produce a valid review of the topic. The words and/or phrases used in the search were: hydrotherapy, Halliwick, Ai Chi, Watsu, Bad Ragaz, chronic obstructive pulmonary disease (COPD), pulmonary rehabilitation, emphysema, chronic bronchitis and the anti-inflammatory effects of exercise. After relevant studies were sourced through these databases the reference lists of these articles were perused to identify further references relating to the topic and these were then sourced through the relevant journal databases. This process was carried out until no further articles were identified for analysis. A hand search for any journal articles which could not be obtained electronically was conducted. An inter-library loan was requested when needed to obtain related articles.
3.4 INCLUSION CRITERIA
Articles in the form of systematic reviews, randomised controlled trials, clinical trials and case studies pertaining to adult patients with non-acute, non-terminal COPD and the use of hydrotherapy were included in this review.

3.5 EXCLUSION CRITERIA
No articles based on animal studies were included. No articles on Kneipp therapy were included as hydrotherapy in this format does not include the immersion of the patient in water.

3.6 DATA COLLATION
Once the material was collected it was read and assessed for inclusion according to its type by two independent reviewers (DS and HvA). A hierarchy of quality of articles was established and the articles were evaluated using the LOW critical evaluation tool (Lewis, Williams and Olds, 2007). Although the LOW tool is mainly structured to evaluate RCTs it was chosen as it provided a comprehensive basis for the evaluation of the non-randomised trials in the studies under review and allowed for them to be accurately appraised and scored with a uniform methodology. The data were compared and similarities and/or differences in results were noted. The results obtained were reported and the effectiveness of the treatment was assessed by looking at cardiopulmonary physiological outcomes as well as general outcomes. The review was a descriptive study but contains simple statistical data (in the form of tables) to record the quality of the material available for review and to afford correlation of the results and conclusions obtained from the information under review.

3.7 DATA ANALYSIS
An initial visual data analysis was undertaken to assess the availability, quality and relevance of data in the literature. As this was a qualitative analysis most reporting was of a descriptive nature. However, some results were tabulated to clarify the data.

3.8 ETHICAL CONSIDERATIONS
Ethical clearance for this study was obtained from the Human Research Ethics Committee of the University of the Witwatersrand, Johannesburg, South Africa (see Appendix B). Under no circumstances did this researcher copy or plagiarise another person’s work. This researcher observed all copyright instructions. If there was a need to reproduce a table or figure from a study under review, the researcher undertook to contact the authors or publishers of that article and made a formal request, in writing, for permission to reproduce the item. The source was then cited in the text. All articles, reports, books, any other form
of publication and web-sites used were cited within the text and within a comprehensive reference list at the end of the dissertation. Permission for the change of title of this dissertation was obtained from the appropriate authority (Appendix C).
CHAPTER 4

4. RESULTS
In order to assess the results in a logical, orderly and appropriate manner the guidelines put forward by Pai (2001), in the Cochrane Reviewer’s Handbook 4.1.4 (October 2001) and the systematic review format used by Lewis, Williams and Olds (2007) were followed.

4.1 IDENTIFICATION AND SELECTION OF STUDIES
After a thorough search of the specified databases and the perusal of reference lists on sourced articles, it was found that there was a very limited number of studies, in the available literature, which investigated outcomes related to an intervention using a population of patients with COPD and in which the intervention included an aquatic component.

Relevant articles were identified using different keywords or phrases within the same or different data bases. In some instances studies were duplicated, in different search engines, using the same input criteria (Table 4.1). No studies relating to any of the phrases or keywords were found in the Cochrane review data base or the CINAHL, MEDLINE or SPORTSDiscus databases.

Thirty-seven studies were retrieved during the search. Nine studies met the inclusion criteria. All of these studies were published in full between 1996 and 2009. One of the included studies, translated from Portuguese, had an English abstract. The total number of patients involved in the included studies (n = 9) was 207 and the sample sizes in the individual studies varied between 10 and 43 participants.
Table 4.1: The Allocation of Studies According to Search Engines and Keywords or Phrases

<table>
<thead>
<tr>
<th>Keywords</th>
<th>Search engine</th>
<th>Number of studies retrieved</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrotherapy and chronic obstructive pulmonary disease</td>
<td>Pubmed</td>
<td>5</td>
<td>Four of these studies were sourced through Pubmed</td>
</tr>
<tr>
<td></td>
<td>Scopus</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Google scholar</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Ai Chi and chronic obstructive pulmonary disease</td>
<td>All search engines</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Halliwick and chronic obstructive pulmonary disease</td>
<td>All search engines</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Watsu and chronic obstructive pulmonary disease</td>
<td>All search engines</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Bad Ragaz and chronic obstructive pulmonary disease</td>
<td>All search engines</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Pulmonary rehabilitation and hydrotherapy</td>
<td>Pubmed</td>
<td>1</td>
<td>These two studies were the same</td>
</tr>
<tr>
<td></td>
<td>Scopus</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Google scholar</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Chronic bronchitis and hydrotherapy</td>
<td>Pubmed</td>
<td>2</td>
<td>This study has already been sourced in Pubmed using the keywords</td>
</tr>
<tr>
<td></td>
<td>Scopus</td>
<td>0</td>
<td>hydrotherapy in chronic obstructive pulmonary disease.</td>
</tr>
<tr>
<td></td>
<td>Google scholar</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Emphysema and hydrotherapy</td>
<td>Pubmed</td>
<td>3</td>
<td>These two studies have been sourced through Pubmed</td>
</tr>
<tr>
<td></td>
<td>Scopus</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Google scholar</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Anti-inflammatory effects and hydrotherapy</td>
<td>Pubmed</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scopus</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Google scholar</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Of the 37 studies sourced 28 did not meet the inclusion criteria for this review (Figure 4.1). Studies were excluded as follows: a) four of the studies fell outside the required chronological period of 1996 to 2009 [two studies on the topic of chronic bronchitis and hydrotherapy were dated 1973 and 1974 respectively, a study on emphysema and hydrotherapy was dated 1965 and was in German with no abstract available, a study on hydrotherapy and COPD was excluded as it was published in 1976]; b) a study on hydrotherapy and COPD did not meet the inclusion criteria as it had no abstract and was written in Japanese; c) although a total of 22 studies were sourced in relation to the anti-
inflammatory effects of hydrotherapy none of the studies dealt with the treatment of COPD and were therefore excluded.

Figure 4.1: Study Selection Process

4.2 DESCRIPTION OF THE STUDIES
The focus and emphasis of this review is to assess the volume of evidence available and qualitatively review the literature on the role of hydrotherapy in the management of COPD. Using the inclusion and exclusion criteria, previously outlined, out of a total of 37 studies which were sourced, nine were considered to offer a resultant contribution to the topic under review. Due to the limited number of studies on the subject the type of study for inclusion could not be limited to randomised controlled trials only.

4.2.1 Volume of Evidence According to Lloyd-Smith (1997) Hierarchy
The nine studies which were included in the systematic review were assessed using the Lloyd-Smith (1997) hierarchy of evidence in order to evaluate the type of study under examination.
Table 4.2:  Hierarchy of Evidence Ranking System (Lloyd-Smith, 1997)

<table>
<thead>
<tr>
<th>Level</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>Meta analysis of randomised controlled trials</td>
</tr>
<tr>
<td>1b</td>
<td>One individual randomised controlled trial</td>
</tr>
<tr>
<td>2a</td>
<td>One well-designed, non-randomised controlled study</td>
</tr>
<tr>
<td>2b</td>
<td>Well-designed quasi-experimental study</td>
</tr>
<tr>
<td>3</td>
<td>Non-experimental descriptive studies-comparitive/case studies</td>
</tr>
<tr>
<td>4</td>
<td>Respectable opinion</td>
</tr>
</tbody>
</table>

According to this methodology the studies fell into the following categories:

There were no studies which fell into the highest category of 1a. Two studies were allocated to the 1b category (Wadell et al, 2004; Wadell et al, 2005). Two studies were classified as 2a (Kurabayashi et al, 1998; Kurabayashi et al, 2000) and one study was classified as 2b (Severino, Morano and de Sousa Pinto, 2007). The remaining four studies fell into the level 3 grouping (Perk, Perk and Bodén 1996; Kurabayashi et al, 1997; Kurabayashi, Machida and Kubota, 1998; Rae and White, 2009).
Figure 4.2: The LOW (Lewis, Olds and Williams) Critical Appraisal Tool
Reproduced with the kind permission of L Lewis (Lewis, Williams and Olds, 2007)
The LOW critical appraisal tool was used to evaluate the nine studies in this qualitative systematic review (Figure 4.2). All the included studies were similarly assessed and scored using the LOW format. On the advice of L. Lewis, of the LOW critical appraisal team, the critical appraisal scoring protocol applied was to answer the questions in each category and to apply a majority vote to the appraisal thereof. As is required in a valid systematic review two independent reviewers (DS and HvA) evaluated the available literature and discussion was held, where necessary, until consensus was reached regarding the quality of each study under review.

Table 4.3: Methodological Quality of the Included Studies (n = 9) Based on the LOW Critical Appraisal Tool

<table>
<thead>
<tr>
<th>Study</th>
<th>LOW Critical Appraisal Rating</th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perk, Perk and Bodén, 1996</td>
<td>yes yes yes yes yes</td>
<td>5</td>
</tr>
<tr>
<td>Kurabayashi et al, 1997</td>
<td>yes</td>
<td>2</td>
</tr>
<tr>
<td>Kurabayashi, Machida and Kubota, 1998</td>
<td>yes</td>
<td>1</td>
</tr>
<tr>
<td>Kurabayashi et al, 1998</td>
<td>yes yes yes</td>
<td>2</td>
</tr>
<tr>
<td>Kurabayashi et al, 2000</td>
<td>yes yes yes yes yes yes</td>
<td>5</td>
</tr>
<tr>
<td>Wadell et al, 2004</td>
<td>yes yes yes yes yes yes yes</td>
<td>8</td>
</tr>
<tr>
<td>Wadell et al, 2005</td>
<td>yes yes yes yes yes yes</td>
<td>7</td>
</tr>
<tr>
<td>Severino Morano and de Sousa Pinto, 2007</td>
<td>yes yes yes yes yes</td>
<td>4</td>
</tr>
<tr>
<td>Rae and White, 2009</td>
<td>yes yes yes yes yes</td>
<td>6</td>
</tr>
</tbody>
</table>

As can be seen in table 4.3 the methodological quality of the included studies ranged from one point (Kurabayashi, Machida and Kubota, 1998) to eight points (Wadell et al, 2004). However, it must be noted that only studies with a separate control group would be able to achieve a score of nine. Eligibility criteria, so that patient recruitment could be repeated, were not stated in five studies (Kurabayashi et al, 1997; Kurabayashi, Machida and Kubota, 1998; Kurabayashi et al, 1998; Kurabayashi et al, 2000; Severino, Morano and de Sousa Pinto, 2007). Four studies had no exclusion criteria stipulated (Kurabayashi et al, 1997; Kurabayashi, Machida and Kubota, 1998; Kurabayashi et al, 1998; Kurabayashi et al, 2000) and one had no clearly defined inclusion criteria (Severino, Morano and de Sousa Pinto, 2007).
The recruitment process was likely to produce bias in regard to the patients’ positive or negative response to the intervention in five of the studies (Perk, Perk and Bodén; Kurabayashi et al, 1997; Kurabayashi, Machida and Kubota, 1998; Kurabayashi et al, 1998; Severino, Morano and de Sousa Pinto, 2007).

A power calculation and justification for the sample size was given in only two of the included studies (Wadell et al, 2004; Wadell et al, 2005). These two studies also were the only studies to include a separate control group.

The stability of baseline measurements and randomisation of interventions were not clearly stated in four studies (Perk, Perk and Bodén, 1997; Kurabayashi et al, 1997, Kurabayashi, Machida and Kubota, 1998; Kurabayashi et al, 1998). In the study by Kurabayashi et al (2000) the number of patients randomly assigned to each group and the order of interventions were not made clear. Whether the interventions were executed on the same day and the time of day they were done was also unclear. Outcome measurements were not measured accurately and therefore bias could not be minimised in two studies (Kurabayashi et al, 1997; Kurabayashi, Machida and Kubota, 1998). In the study by Perk, Perk and Bodén (1996) no clear indication was given by the authors as to the frequency of the intervention, the number of times the intervention occurred and the total duration of the intervention.

No studies mentioned potential confounding factors or the possible impact of confounding factors on outcomes measured. However, two studies did mention that the sample size used was small (Wadell et al, 2005 and Severino, Morano and de Sousa Pinto, 2007).

In the study by Wadell et al (2004) the researchers calculated the effect size (ES) for the ISWT, ESWT and SF-36. Otherwise no confidence interval data was given in any of the studies. Attrition/withdrawal data were not given in four studies (Kurabayashi, Machida and Kubota, 1998; Kurabayashi et al, 1998; Kurabayashi et al, 2000; Severino, Morano and de Sousa Pinto, 2007).

The tools used for the calculation of QoL and ADL in two studies (Kurabayashi et al, 1997; Kurabayashi, Machida and Kubota, 1998) were possibly not appropriate for the measurements of these two parameters. In the study by Kurabayashi et al (1998) no analysis of data/variance (ANOVA) calculation was mentioned in the comparison of the two groups.
Overall there was a direct correlation between the LOW critical appraisal score rating and the methodological quality of the studies. A small LOW score was a valuable indicator of the possible poor validity of the results and a high LOW score indicated a study with possible valid outcomes.

The comparative analyses of each study by the two reviewers (DS and HvA), using the LOW critical appraisal tool, is shown in Appendix D. The two reviewers discussed points of difference in the scores and the final results of the LOW critical appraisal, of the included studies, is recorded in Table 4.3.

4.2.2 Generalisability of Evidence

All nine of the studies (100%) looked at the cardiorespiratory effects of hydrotherapy in the rehabilitative management of patients with stable COPD. As this was a review of hydrotherapy interventions the temperature of the water and the depth of immersion of the patients was an important consideration. Four studies (44%) used hot spring water at a temperature of 38°C; all other studies (56%) were done in thermoneutral water (29°C-34°C). All studies (100%) used the head-out methodology of immersion with the water level up to shoulder level.

Therefore the generalisability of the evidence was classified as good as all researchers focussed on the rehabilitation of patients with COPD immersed to the same depth in water of a similar temperature.

4.2.3 Participants

The majority of the patients who participated in the studies were elderly and between 60 and 80 years of age. However, in one study (11%) the age variable spanned 45 years (35-80 years).

Of the nine included studies (100%) which investigated the effects of a hydrotherapy intervention in the COPD population, two studies (22%) included both asthmatic patients and those with pulmonary emphysema categorising the patients as uniformly suffering from COPD. Two studies (22%) by Kurabayashi, Machida and Kubota (1998) and Kurabayashi et al (2000) stipulated that the patients with COPD were suffering from pulmonary emphysema and the other five studies (56%) did not categorise the patients into a COPD subgroup. All the patients in the included studies had stable COPD. One study (11%) by Rae and White (2009) defined the patients’ COPD level as mild or moderate. Four studies (44%) did not stipulate the patients’ COPD level (Kurabayashi et al, 1997; Kurabayashi, Machida and Kubota, 1998; Kurabayashi et al, 1998; Kurabayashi et al, 2000). Calculation
of the COPD levels of the patients in the study by Perk, Perk and Bodén (11%) was possible from the FEV₁ predicted levels given. The COPD level of the patients in this study was moderate to severe. Severino, Morano and de Sousa Pinto (2007), in their study (11%), stipulated that the patients had moderate to severe COPD. Only two studies (22%) (Wadell et al, 2004; Wadell et al, 2005) used the GOLD criteria, which stipulates 50%≤FEV₁< 80% predicted for moderate COPD and 30%≤FEV₁<50% predicted for severe COPD, as a definitive measure to report that the patients in their studies fell into the moderate to severe COPD category. These were the only studies (22%) that had a separate control group (Wadell et al, 2004; Wadell et al, 2005).

An included study (Kurabayashi, Machida and Kubota, 1998), which measured the effect of hydrotherapy on EF, included four patients with the New York Heart Association (NYHA) functional class I, six patients with NYHA functional class 2 and two patients with NYHA functional class 3 (Kurabayashi, Machida and Kubota, 1998). [NYHA functional class 1 is representative of mild heart failure with no limitation of physical activity; NYHA functional class 2 represents mild heart failure with slight limitation of physical activity and NYHA functional class 3 where there is moderate heart failure with marked limitation of physical activity].

4.2.4 Intervention

Four of the included studies (Kurabayashi et al, 1997; Kurabayashi, Machida and Kubota, 1998; Kurabayashi, et al, 1998; Kurabayashi et al, 2000) looked at various applications of breathing exercises performed in a hot spring water pool. In two of the studies (Kurabayashi et al, 1997; Kurabayashi, Machida and Kubota, 1998) the breathing exercises were done with the patient breathing out into the water through the mouth while the nose was submerged 3-5 centimetres below the water surface. The duration of the session was 30 minutes and the procedure was done six times a week for a period of two months. One of the other two studies (Kurabayashi et al, 1998), applying the same breathing regime, looked at three different breathing protocols (A, B and C) where the total exercise period per week in the first intervention (A) was 20 minutes per week (10min/day, 2days/week); the second intervention (B) was 120 minutes per week (20min×2/day, at 10:00am and 3:00pm, 3 days /week) and the third (C) was for 120 minutes per week (20min/day, 6days/week) (Kurabayashi et al, 1998). In the remaining study the breathing exercises were carried out in the water but one group breathed out into air while the other breathed out into the water (Kurabayashi et al, 2000).

All the other included studies looked at varied PR programmes. Four of these studies looked at the same exercises being performed on land versus in water (Perk, Perk and
Bodén 1996; Wadell et al, 2004; Wadell et al, 2005; Severino, Morano and de Sousa Pinto, 2007) and the fifth study was based entirely in a swimming pool (Rae and White, 2009). In the two studies by Wadell et al (2004) and Wadell et al (2005) there was a high intensity physical training programme incorporating 45 minute sessions three times a week for a total period of 12 weeks. The exercises in these studies had strength and endurance components and exercised the whole body (Wadell et al, 2004; Wadell et al, 2005). The study by Perk, Perk and Bodén (1996) involved three periods of submaximal dynamic arm and upper body exercise each lasting for three minutes followed by a two minute rest period for a total time of 15 minutes on land and 15 minutes in water. The land-based 15 minute exercise session preceded the water-based session and there was a 10-15 minute rest period between the two sessions. It appears that this intervention was only performed once for a total duration of 30 minutes of exercise. Patients performed hydrotherapy once a week together with a land-based rehabilitation exercise programme twice per week in a study by Severino, Morano and de Sousa Pinto (2007). The exercises performed were from the Kabett diagonal weight and place system and the exercise programme lasted for three months. Rae and White (2009) assessed a hydrotherapy exercise programme incorporating 12 sessions over a period of six weeks which consisted of upper and lower limb exercises. Education, by a multi-disciplinary team, about various matters relating to COPD, was included as part of the intervention by Severino, Morano and de Sousa Pinto (2007) and Rae and White (2009).

4.2.5 Outcome Measures Assessed
A total of 35 outcome measures were used in the nine included studies and were grouped into the following categories: Cardiovascular; pulmonary; muscle strength/weakness, endurance and power; QoL and ADL outcomes.

The most commonly reported outcome was FEV₁% (n = 6) with two of these measurements being calculated at baseline. (Forced vital capacity (FVC), PaCO₂, PaO₂ were used as outcome measures in four studies (n=4). Forced expiratory volume in one second, maximal expiratory flow at 25% (V₂₅), PF, the Borg dyspnoea scale and ISWT were reported outcome measures in three studies (n = 3).

4.2.5.1 The Cardiovascular Outcomes
In a study, in which dynamic arm exercises were done on land for a 15 minute duration and after a 10-15 minute rest period were repeated in water for a period of 15 minutes, Perk, Perk and Bodén (1996) reported that the resting HR after training in water returned to pretraining levels more quickly than the resting HR after training on land. In an exercise programme using Kabett's diagonal weight and place exercises for the upper limb in which
Group 1 did hydrotherapy and a land-based PRP and Group 2 only did a land-based PRP programme, the researchers found that there was a significant decrease in HR in Group 1 ($p< 0.01$) and Group 2 ($p< 0.05$). The decrease in HR was greater in Group 1. HR measurements were taken during the time that the weights were being changed (increased). The programme ran for a period of three months with the patients exercising three times a week. Group 1 exercised once a week in water and twice a week on land (Severino, Morano and de Sousa Pinto, 2007). The mean HR, during training sessions, was shown by Wadell et al (2004) as a percentage of peak heart rate and overall was found to be lower in patients exercising in water for 12 weeks. However, during the cycle ergometer training in the water, a small but significant increase in peak heat rate ($H_{\text{peak}}$) was recorded (Wadell et al, 2004).

Perk, Perk and Bodén (1996) reported a decrease in SBP of 14mmHg and a decrease in DBP of 6mmHg when the patients with COPD were immersed in water compared to the resting values on land. The researchers pointed out that, because the land-based exercise programme and the subsequent 10-15 minute rest period preceded the water–based blood pressure readings, the method employed might have some impact on recorded blood pressure readings. The blood pressure values were given as mean plus standard deviation (+SD) but it was not clear in this study as to whether it was a singular intervention or whether the intervention took place over a number of times. No information regarding the number of times the measurements were taken or intervals between measurements or total duration of the intervention were stipulated by the researchers.

Ejection fraction was the primary outcome measured in a study by Kurabayashi, Machida and Kubota (1998) in which the intervention involved a programme of breathing exercises in water for 30 minutes per day, six days a week over a period of two months. The researchers reported that the EF was significantly increased after each 30 minute exercise session ($p<0.01$) and was also significantly increased at the end of the two month programme ($p<0.05$). In a comparison of three breathing exercise protocols, undertaken in water, by Kurabayashi et al (1998). EF was increased significantly in all three groups 10 minutes after exercise on the first day ($p< 0.01$) but there was no increase reported after 24 hours after the exercise session on the first day. The longer intervention periods of 120 minutes per week in protocols B and C proved to result in a significant improvement in EF at the end of the study period (after two months).

Kurabayashi, Machida and Kubota (1998) measured $L_{\text{vd}}$ and $L_{\text{vs}}$ dimensions and found that the $L_{\text{vd}}$ and $L_{\text{vs}}$ dimensions, at rest, decreased significantly after the two month exercise programme of breathing exercises in water ($p< 0.01$).
4.2.5.2 The Pulmonary Outcomes

Many pulmonary outcomes were measured across the nine included studies. Six studies (n=6) reported on the predicted FEV₁% but in two studies these measurements were recorded in baseline data only. Four of the included studies (n=4) listed outcome data for predicted FVC, Pa CO₂ and PaO₂. Three studies (n=3) reported results for V₂₅, FEV₁, PF and Borg ratings for dyspnoea (one only measured this outcome at baseline). The ratio of FEV₁/FVC, SaO₂, VO₂peak and RR were reported in two of the included studies (n=2). Rate pressure product (RPP), peak expiratory flow (PEF), maximum expiratory flow at 50 % (V₅₀), VC, CRQ-SR rating for dyspnoea, base excess (BE) and the acid/alkaline level (pH) of arterial blood were each only recorded once (n=1) within two studies.

The two included studies which looked at breathing exercises in water where the patients exhaled into the water for an exercise period of 30 minutes a day /six days a week for two months both reported a significant increase in FEV₁% (p<0.37 and p< 0.05 respectively) (Kurabayashi et al, 1997; Kurabayashi, Machida and Kubota, 1998). These two studies reported no significant outcomes relating to FVC% but both recorded a significant decrease in PaCO₂ levels with no change in PaO₂ levels after two months of hydrotherapy.

In the study by Kurabayashi et al (1998), where three hydrotherapy breathing protocols were tested, a significant increase in FEV₁% (p< 0.01) was recorded in protocols B and C. No significant change was noted in FVC% in any of the three protocols. The PaCO₂ levels were significantly decreased with protocol B (p< 0.05) and protocol C (p<0.05). However, the PaO₂ levels were increased with protocol B only (Kurabayashi et al, 1998).

Kurabayashi et al (2000) reported on a study where patients participated in a breathing exercise programme in water. Group A exhaled into the water while Group B, similarly immersed in the water, exhaled into air. FEV₁% was significantly increased in the patients in Group A (p< 0.018). There was also an inclination towards an increase in FVC% in patients in Group A (p< 0.058) and significant increase in PaO₂ levels (p< 0.010) and a decrease in PaCO₂ (p<0.040) in this group. Group B showed no change in PaCO₂ or PaO₂ levels (Kurabayashi et al, 2000).

Peak flow reportedly showed a tendency to increase in a study of breathing exercises in water where the patients with COPD exhaled into the water (Kurabayashi et al 1997). It was found to increase in patients with COPD in Group A, who exhaled into water as opposed to Group B who exhaled into air in a study by Kurabayashi et al (2000) and was increased significantly in protocols B and C (p< 0.05) in a study by Kurabayashi et al (1998).
Maximum expiratory flow at 25% was not found to change in three of the included studies (Kurabayashi et al, 1997; Kurabayashi et al, 1998; Kurabayashi et al, 2000). In one of the included studies maximum expiratory flow at 50% (V_{50}) and the ratio of V_{50}/V_{25} also remained unchanged (Kurabayashi et al, 1997).

Dyspnoea ratings, according to the Borg scale, were found to be higher both in the horizontal and the vertical pulling exercises in water (p<0.01) and this was more marked in the patients who had a poorer lung function (Perk, Perk and Bodén, 1996). Ratings for dyspnoea averaged a score of four on the 0-10 Borg scale in both training groups (on land or in water) during the physical group training sessions performed in a study by Wadell et al (2004). The ratings for perceived exertion (6-20 Borg scale) averaged 14 (Wadell et al, 2004). In the study by Severino, Morano and de Sousa Pinto (2007) the patients in the water group seemed to be able to have a greater tolerance to the loads imposed on them before complaining of dyspnoea and they also rated the degree of their dyspnoea lower than those in the land-based group. After a swimming pool-based PRP of 12 exercise sessions over a period of six weeks significant improvements in the dyspnoea scores, according to the CRQ-SR, were seen in the participants (Rae and White, 2009).(No SD statistical data were reported).

Only one of the included studies looked at BE and the pH of arterial blood. The researchers found no change to these outcomes compared to baseline after the programme of breathing exercises conducted in water over a period of two months (Kurabayashi et al, 1997). Although during a cycle ergometry test, the VO_{2peak} was increased in all three groups in a study looking at group physical training in water and on land (including a control group with no intervention) a significant increase was seen in the control group (p<0.018) and in the water group (p< 0.008). However, the researchers state that it is questionable whether the improvement in the control group is clinically relevant (Wadell et al, 2004).

There was a significant improvement in SaO_{2} levels (p< 0.008) in the group that underwent training once a week in a pool together with two land-based exercise sessions compared to the group which only underwent land –based exercise three times/ week (Severino, Morano and de Sousa Pinto, 2007). SaO_{2} levels exhibited by the water group showed a 33.3% improvement over those recorded for the land based PRP group, after the 12 week programme. The oxygen saturation showed a significant 2% (p< 0.05) fall during the vertical pulling arm exercises conducted in water in the study by Perk, Perk and Bodén (1996). The other dynamic arm exercises had no effect on the SaO_{2} levels in the outcomes of this particular study. In the same study it was reported that, with immersion in the water, VC was significantly decreased by 12% (p< 0.001), FEV_{1} significantly decreased by 14%
(p<0.001), PEF significantly decreased by 18% (p< 0.001) but the ratio FEV₁/VC remained the same. Very little difference, compared to baseline measurements, was seen in FEV₁ and FEV₁/ FVC in the swimming pool-based study by Rae and White (2009) (No SD data were given for these parameters).

The RR decreased significantly (p< 0.046) after a three month aquatic exercise intervention in the study by Severino, Morano and de Sousa Pinto (2007). However, Perk, Perk and Bodén (1996) contrarily noted that, during some of the exercises performed in water in their study, breathing required more effort. Both of these studies looked at exercises performed with the upper limbs in an aquatic environment. Perk, Perk and Bodén (1996) used a 15 minute regimen of submaximal dynamic arm and upper body exercises (three exercises performed for three minutes each with a two minute rest period between each exercise), firstly performed on land and then in water for a total exercise time of 30 minutes per session. The exercise regimen included horizontal (rowing action and curtain closing action) and vertical (bucket lifting action) weight pulling exercises. Severino, Morano and de Sousa Pinto’s (2007) study used Kabett’s diagonal weight and place system and was carried out once a week in the water and once a week on land in the water group and twice a week on land in the land group. The programme was continued for 12 weeks. Perhaps the improved physical conditioning and fitness of the patients due to the length of the intervention in the study by Severino, Morano and de Sousa Pinto (2007), accounted for the reported beneficial outcomes. The total duration of the intervention by Perk, Perk and Bodén (1996) was not clearly stated and so it is difficult to assess whether it was of shorter duration than that of Severino, Morano and de Sousa Pinto (2007) and whether physical conditioning or lack thereof was reflected in the response seen in the patients with COPD in the study by Perk, Perk and Bodén.

4.2.5.3 Muscle Strength/Weakness, Endurance and Power

Wadell et al (2004) conducted a trial in which patients with COPD performed high intensity physical training in water or on land for 45 minutes per session, three times a week for a period of 12 weeks. These researchers reported an increase in walking distance for both groups. In the ISWT (the result was significant in the land group) and in the ESWT (the result was significant in the water group). The mean effect size (ES) values with 95% confidence intervals were given for these results. During the cycle ergometer tests in the same study both the water group and the land group increased the time cycled. Mean workload was also increased significantly across all three groups with an increment of 6.7, 9.3 and 10.0 W for control, water and land groups respectively. Only baseline data relating to between-group comparisons for ISWT and ESWT in the land group, water group and control group were reported by Wadell et al (2005) and these data were not significant.
After an aquatic intervention of 12 sessions of exercise over six weeks in a swimming pool, Rae and White (2009) reported that ISWT scores improved significantly.

In a study by Wadell et al (2005), the focus of the research was on the effects of a three-month physical training programme on muscle performance in patients with COPD. Maximal dynamic strength and endurance of the thigh muscle was measured pre and post training. The patients that were randomised to the land-based intervention group improved in both maximal dynamic extension by 16% (p< 0.009) and in maximal dynamic flexion by 27% (p< 0.006). Improvements in maximal dynamic flexion of 40% (p< 0.001) and 7% (p<0.016) were noted in the water-based intervention group and the control group respectively.

In the water-based intervention group, there was an improvement of 8% in maximal dynamic extension (p<0.023). The endurance test resulted in no significant change observed in any of the groups. The researchers also noted a high non-responder grouping with regard to the ability to perform 100 knee extensions at baseline. They noted some degree of correlation between lower body mass index (BMI) and better strength in knee extension (Wadell et al, 2005). They concluded that physical training on land or in water has beneficial outcomes regarding maximal thigh muscle strength in patients with COPD.

Severino, Morano and de Sousa Pinto (2007), using the Borg tool, found that upper limb fatigue changed little between the group that performed the exercise in water as well as a land-based PRP compared to the group who only performed the land-based PRP. The study by Perk, Perk and Bodén (1996) was the only one to record a rating of effort using the Borg scale for this parameter. These researchers found that during exercise 1, horizontal weight pulling as in a rowing motion, a resultant significant increase in this rating from 12-14 (p<0.01) was recorded. Wadell et al (2004) reported that the physical component score (PCS) of the SF-36 showed that the water intervention produced a significant improvement (p=0.015) in physical health. This study demonstrated that the mean workload was improved in the water group compared to that in the control group.

With regard to maximal load, Severino, Morano and de Sousa Pinto (2007) reported that the patients in the group which had the hydrotherapy component in their regimen performed better than those in the land-based group. They qualified this observation by saying that those patients were able to withstand higher loads during the intervention, showing a greater physical fitness.

4.2.5.4 Quality of Life (QoL) and Activities of Daily Living (ADL) Outcomes

Wadell et al (2004), using the SGRQ, found that the patients in the water group showed a small significant improvement in total score but also showed a significant decrease in activity score where the decrease was by 5.1 points and four units was regarded as the
threshold for clinically significant change. This was regarded as significantly improved compared to the other two groups (p=0.009). The control group showed deterioration according to the SGRQ rating system which was reported to be significant (p=0.034). According to the generic SF-36 questionnaire, Wadell et al (2004) found that the PCS for their patients were lower than those recorded for the general Swedish population over 65 years of age. The mental component scores (MCS) were compatible. However, an increase of six points in the PCS for the water group post intervention was considered to be a clinically important difference and this change was significant compared to the other two groups. The ES achieved in their study for exercise capacity and HRQoL were between 0 and 0.7 in the water group, between 0 and 0.3 in the land group and between 0 and -0.4 in the control group. In the study by Kurabayashi et al (1997) the modified version of the Flanagan (Flanagan, 1982) classification was used to assess changes experienced by patients with regard to QoL. Of the 22 patients in the study five showed objective signs of improvement in their QoL. Kurabayashi et al (1997) used the Hugh-Jones (Hugh-Jones, 1952) classification to assess functional performance. The authors recorded that five of the patients demonstrated objective improvement on this scale. Kurabayashi, Machida and Kubota (1998), using the same classification tool, found that after discharge five patients showed an improved QoL. Rae and White (2009) used a subjective evaluation of the treatment by the patients with COPD. Qualitative interviews, conducted after the intervention, were audio-taped, transcribed and analysed. Patients evaluated the programme as good or very good. According to the emotional scores and mastery scores before and after the intervention as measured by the CRQ-SR tool there was an improvement in both these parameters. However, the researchers did not give SD information for these data.

The calculation of ES is a reliable method to assess the degree to which a specific intervention influenced the outcomes measured. The LOW critical appraisal tool (Question 9a) was used to determine whether the ES was presented or could be calculated from the data presented in the included studies. Although five of the included studies reported on mean and SD data for outcomes assessed as well as the size of the p value (Perk, Perk and Bodén, 1996; Kurabayashi et al, 1997; Kurabayashi et al, 1998; Wadell et al, 2004; Severinol, Morano and de Sousa Pinto 2007) it was not possible to assess ES as these authors did not assess similar outcome measures and there were no control groups in their study designs except for Wadell et al (2004). Wadell et al (2004) had reported on ES for the ISWT, ESWT, MCS and PCS. It was also not possible to assess ES for the semi-randomised RCT conducted by Wadell et al (2005) as the authors presented their data in the form of medians and interquartile ranges and not means and SD.
CHAPTER 5

5. DISCUSSION

This qualitative systematic review looked at the specific population of patients with COPD and the effect of hydrotherapy management of their disease. It examined how hydrotherapy influenced cardiopulmonary outcomes, endurance, the patient’s ability to perform ADL and the HRQoL in patients with COPD. It also tried to establish whether there was any evidence connecting exercise in water to an anti-inflammatory effect similar to that seen in land-based exercise.

The fact that only nine studies met the inclusion criteria for this qualitative systematic review attested to the fact that there were very few studies which assessed the role of hydrotherapy in the management of COPD. The range of scores that the studies obtained on the “LOW” critical appraisal rating tool showed that the included studies were of varied quality. Therefore, in the interpretation of the results and conclusions drawn from these studies, the LOW critical appraisal scores were taken into account when considering the possible validity of the researcher’s findings. The limited information on the topic, the diversity of the interventions and very few related measured outcomes also made accurate assessment of the validity of the results difficult.

The cardiopulmonary outcomes were examined in relation to a number of interventions. The included studies (n=4), which examined the role of breathing exercises being undertaken in hot spring water (38°C), produced some interesting results. It was noted that the cardiovascular parameters showed positive increases when these were examined in the three studies involving immersion with inhalation in air followed by exhalation into water. In relation to the pulmonary parameters the pulmonary measurements showed positive and beneficial outcomes for the patients with COPD when immersion was linked to the act of exhaling into the hot spring water. In particular, an improvement in the blood gas results (PaCO₂ decreased and PaO₂ increased) was noted when the patients breathed out into the water. When the patients were instructed to breathe out into air, while immersed in the water, no improvement in these parameters was noted and this appears to indicate that the exercise of exhaling into the water had the beneficial effects seen in the results obtained. A possible explanation for this finding is that exhalation against the pressure of the water would create a positive expiratory pressure in the airways. This positive expiratory pressure would open partially collapsed airways as it assists with secretion clearance through collateral ventilation pathways between the alveoli and bronchioles. This would create a greater surface area for gas exchange, hence the improvement in ventilation observed (Pryor and Prasad 2008, 151). Expiration against resistance enforces more work on the
inspiratory muscles to overcome this resistance and therefore increases lung volumes which would lead to improved ventilation (Fagevik-Olsen, Lonroth and Bake 1999; Lumb 2005, 49-50).

It was demonstrated that, when such a breathing exercise regime, in water, had a duration of 120 minutes/week (Kurabayashi et al, 1998) or 150 minutes /week (Kurabayashi, Machida and Kubota, 1998), a significant improvement was achieved in cardiopulmonary outcomes when compared to those achieved in an intervention of only 20 minutes/week (Kurabayashi et al, 1998). However, it was also demonstrated that PaCO$_2$ decreased most when the exercise programme was performed once a day for six days a week as opposed to twice a day for three days per week. Sessions were of 20 minutes duration in both scenarios (Kurabayashi et al, 1998). Partial pressure of oxygen in arterial blood was only seen to increase significantly ($p<0.01$) in protocol B of the study, which consisted of breathing out into water for 20 minutes, twice a day (once in the morning and once in the afternoon), for three sessions a week over a period of two months (Kurabayashi, Machida and Kubota,1998).

Hence, it appears that physiotherapy for COPD, in the form of a hydrotherapy breathing regime, may require an intervention of a particular duration and frequency to improve clinical parameters. However, the validity of these findings was questioned by this researcher as the three studies from which this assumption has been made had a poor rating on the LOW critical appraisal scale.

As previously reported in 2.6.1 of this review, Weiner et al (2003a) concluded that when land-based EMT was performed by patients with COPD over a period of three months a significant improvement in exercise performance was noted. Weiner and McConnell (2005) concluded that strength and endurance can be improved when specific training is given to expiratory muscle groups. They also concluded that EMT leads to an increase in exercise performance. The findings in the aquatic breathing exercise regimens reported by Kurabayashi et al (1997); Kurabayashi, Machida and Kubota (1998); Kurabayashi et al (1998) and Kurabayashi et al (2000) have suggested beneficial cardiopulmonary outcomes when patients with COPD make an expiratory muscle effort in order to exhale into water. Therefore, this type of breathing exercise should perhaps be considered as an alternative form of EMT in patients with COPD.

In the cases where the intervention for the patients with COPD was an aquatic physical training programme, either alone or in comparison to a land-based intervention (n=4), a shorter aquatic exercise programme of 15 minutes was not seen to be instrumental in
changing the HR parameter which was measured during the exercise session and also during the 10-15 minute rest period after the exercise session had been completed. However, it was noted by the researchers that post training resting HR levels on land did not return to resting pretraining levels to the same extent as post training resting HR levels in water. Systolic blood pressure and DBP, when measured at rest in water, showed beneficial decreases. The study by Perk, Perk and Bodén (1996) from which these data were taken did not stipulate a total length of time for the duration of the intervention and it seems possible that it was a once-off intervention. In 12 week hydrotherapy interventions the HR was shown to decrease during the training period. It has been documented that immersion up to the xiphoid level (shoulder level) results in a decrease in HR, even without physical activity (Lollgen et al, 1981; Gabrielson, Johansen and Norsk, 1993). A decrease in HR due to immersion alone, which should have been seen in the shorter exercise intervention, was possibly counteracted by an increase in HR brought about by the short, intense exercise intervention thus resulting in an overall stabilisation of the HR.

In general, dyspnoea ratings for the physical training interventions in four of the included studies with both land–based and water-based interventions and one with only an aquatic component (n=5), were low especially in relation to the longer duration interventions. When dynamic upper limb and trunk exercises were done in water for a period of 15 minutes the Borg ratings for dyspnoea were raised and were especially increased in the patients who had poorer lung function. In the 12 week hydrotherapy and PR intervention, where the patients with COPD performed Kabett’s diagonal weight and place exercises with the upper limbs, the perception of dyspnoea was seen to be delayed during the exercise programme in the water and was of a lower rating when assessed. This type of effect was also shown when patients with COPD took part in a swimming pool-based physical training programme for a period of 12 weeks and an average Borg scale rating of four was seen in another physical training intervention over a period of 12 weeks. All three of these advantageous results for the reduction of dyspnoea were seen to be associated with an aquatic intervention once, twice or three times per week over a period of 12 weeks. The frequency of application of the intervention varied; the interventions, although of a physical nature also varied and the only common denominator to all three studies seemed to be the total duration of the intervention period (12 weeks). As the scores for dyspnoea for the 15 minute intervention were raised but for the longer interventions were decreased there may be a case to suggest that hydrotherapy interventions involving physical exercise programmes in patients with COPD could have favourable outcomes in relation to the reduction of dyspnoea if the intervention lasted for at least 12 weeks. In general no detrimental effects of hydrotherapy on cardiopulmonary outcomes were reported.
Dysfunction of the skeletal muscles has been recognised as part of the COPD profile (Casaburi, 2001). The hydrotherapy physical training interventions for patients with COPD included in this review looked to address the weakness of the skeletal musculature. Wadell et al (2005) conducted a three month trial to assess the performance of thigh muscle in patients with COPD undergoing a specific physical exercise programme. They conducted the test on land and in water and found that training in either environment resulted in an improvement in thigh muscle strength but muscle endurance remained at pre-training levels. Increases in work load achieved through cycle ergometry (Wadell et al, 2004), distance walked and endurance on the ISWT (Wadell et al, 2004; Rae and White, 2009) and ESWT (Wadell et al, 2004), duration of time spent on the cycle ergometer (Wadell et al, 2004), maximal dynamic knee flexion and to a lesser extent dynamic knee extension (Wadell, 2005) were reported in relation to hydrotherapy intervention. It may be concluded from these results, although few in number, that aquatic physical exercise training programmes might have some value in raising the functionality and strength of the skeletal muscles in patients with COPD and have an effect on endurance related functions such as walking. Man et al (2009), in a review on exercise and muscle dysfunction in COPD, concluded that an increase in quadriceps muscle strength combined with an increase in endurance was reported by some researchers after a land-based PRP. These authors also reported that a decrease in fatigability of the muscle, a small increase of the fat-free muscle mass and CSA was also found in the studies included in their review. These findings substantiate the reported improvement in quadriceps muscle strength observed by Wadell et al (2005) in relation to the land-based portion of their intervention and may lead to the conclusion that the water-based result may also be feasible. The studies (n=3) which reported these findings were above average according to the LOW critical appraisal rating tool and therefore their findings may be more objectively valid. Therefore, hydrotherapy may be a good alternative form of exercise intervention.

The St George’s respiratory questionnaire results and PCS results (obtained from the SF-36 questionnaire) demonstrated that a hydrotherapy intervention of physical training led to significant improvement in the areas relating to QoL (Wadell et al, 2004). Positive results were demonstrated on the CRQ-SR in the study by Rae and White (2009) where patients with COPD took part in a swimming pool based PR programme of 12 sessions of exercise over six weeks. Although a modified version of the Flanagan classification (Flanagan, 1982) was used to assess QoL in two studies the format for the application of the tool was not clear. The two studies using this tool reported that QoL was improved with the hydrotherapy intervention. However, the studies were of a poor quality as shown by the LOW critical appraisal rating tool. With regard to ADL the studies by Kurabayashi et al (1997) and Kurabayashi, Machida and Kubota (1998) used the Hugh-Jones classification tool (Hugh-
Jones, 1952) where this tool is used to relate the ventilation cost of the performance of a step test to the degree of breathlessness in the patients. It is not related to any specific ADL and therefore the choice of its use in order to assess ADL outcomes is questionable. This classification tool is also of a vintage nature and the Borg classification for dyspnoea is widely used and accepted today. The subjective evaluation of participants’ experience of aquatic therapy was also offered as proof that group hydrotherapy improves QoL and the performance of ADL in patients with COPD (Rae and White, 2009).

The small sample size, the issue of COPD exacerbations and patients’ difficulty with the commute to the swimming pool were raised as factors which influenced the number of patient participating ,at any given time, during the combination aquatic and land-based intervention in the study by Severino, Morano and de Sousa Pinto (2007). Gender related differences in regard to certain outcome measurements were reported by Wadell et al (2005) and Severino, Morano and de Sousa Pinto (2007) but both sets of researchers did not feel that the data were tainted by this fact.

Various authors,[Craig and Dvorak (1968); McArdle et al (1976); Choukroun and Varène (1989) and Choukroun, Kays and Varène (1990)], while conducting trials with healthy individuals, recommended that a water temperature between 29°C and 33°C should be regarded as thermoneutral for the purpose of physical exercise programmes in water. As previously mentioned Israel et al (1989) in a study using healthy individuals reported that, during exercise at approximately 60% of maximum oxygen consumption (VO_{2max}), immersion in 25°C and 21°C water prevented a rise in core temperature. The same exercise regimen in 29°C water led to an increase in core temperature. Hence Israel et al (1989) concluded that, individuals of similar body composition and fitness level, performing cardiovascular exercise, in 21°C or 25°C water temperatures would benefit from the cooling effects resulting from immersion at these temperatures. These researchers also reported that very low intensity hydrotherapy could be performed in 29°C water as they had found that static immersion at that temperature had resulted in a fall in core temperature in the subjects in their study. In the included studies, for this review, no particular reference to water temperature was made in regard to its influence on the outcomes measured. In fact only two of the studies in which physical training programmes were carried out mentioned the exact temperature of the water. Wadell, et al (2004) used a water temperature between 33°C and 34°C and Perk, Perk and Bodén (1996) used water at 32°C. Severino, Morano and de Sousa Pinto (2007) and Rae and White (2009) conducted their interventions in swimming pools where the ambient temperature was probably within the afore-mentioned range. Wadell et al (2005) gave no indication of the water temperature used in their study. However, the four studies included in this review (Kurabayashi et al, 1997; Kurabayashi,
Machida and Kubota, 1998; Kurabayashi et al, 1998; Kurabayashi et al 2000) which looked at various cardiopulmonary outcomes in relation to a hydrotherapy breathing exercise regimen used hot spring water with a temperature of 38°C with no reported ill effects in the patients taking part. Although the breathing exercise interventions were conducted in hot spring water with a temperature of 38°C the exertion factor in these studies was minimal when compared to those involving a physical exercise hydrotherapy programme. Perhaps other factors which should be taken into account in such interventions are that variations in water temperature change the amount of thermal exchange through the skin (Choukroun, Kays and Varène, 1989). Wilmore and Costill (1994) pointed out that the thermal conductivity of water was approximately 26 times that of air and rapid loss of body heat occurred by conduction and convection during immersion. The passage of water over the body during aquatic exercise further accelerates the heat loss (Datta and Tipton, 2006). In this regard it was reported by Kurabayashi et al (1997) that the hot spring water, used in their study, could keep the body warm for longer as materials dissolved in the spring water act as a shield on the skin and prevent evaporation of heat. The gases eminating from the hot spring water could also have a mucolytic effect on the sputum of the patients with COPD and therefore aid with airway clearance. Kurabayashi, Machida and Kubota (1998) cited work done by Tei et al (1995) in which those researchers had found that a bath at 41°C and a sauna bath at 60°C were effective in treating severe heart failure. However, Kurabayashi, Machida and Kubota (1998) reported that those temperatures were too high for 30 minutes of breathing exercises performed by patients with COPD immersed up to shoulder level and so they chose 38°C for their intervention. Kurabayashi, Machida and Kubota reported a significant increase in EF in patients who were immersed in 38°C water but the authors suggested that this improvement was possibly attributable to the hydraulic pressure exerted by the water. By assimilating all the above information this researcher concluded that there is enough proof to suggest that physical training exercise programmes for patients with COPD would be best conducted in thermoneutral water in the 32°C to 34°C range as this appears to be utilised in the COPD population with no adverse effects on the well-being of the patients and beneficial results are also realised in cardiopulmonary outcomes, strength and endurance outcomes, social and psychological areas and ADL and HRQoL. However, the researcher would recommend that for breathing exercise programmes where the patient with COPD is merely standing in the water and exhaling into the water the water temperature used could probably be higher (38°C).
The hydrotherapy PR scenario appears to be well accepted by patients with COPD as many of them are elderly patients and find that the buoyancy of the water assists them to exercise more effectively than on land due to less impact being felt on arthritic joints (Rae and White 2009). Comments relating to social interaction during hydrotherapy and the psychological well-being expressed by patients with COPD, in the aquatic environment, were recorded in a study by Rae and White (2009). Hydrotherapy offers the benefits of low-cost group training and may be seen as an attractive alternative by the patients with COPD who often grow bored with land-based PR (Perk, Perk and Bodén, 1996). The value of this enthusiasm is that the patients are more compliant with hydrotherapy exercise programmes than land-based exercise regimens. It was reported that, six months after the intervention by Rae and White (2009), 50% of the participants were still exercising in the swimming pool twice-weekly. The study by Rae and White (2009) also pinpointed the need for early intervention in the course of the disease in relation to improving the QoL of patients with COPD. These researchers also emphasised the importance of educating the patients with COPD in the early stages of the disease and, in particular, educating them regarding maximising the value of PR.

Based on the information gleaned from this qualitative systematic review, the researcher would like to present possible guidelines for the management of patients with COPD in the hydrotherapy environment:

- Water temperature should preferably be within the range of 27 - 35°C.
- The physiotherapist should be aware of the reported drop in SBP (up to 14mmHg) as well as decrease in DBP (up to 6 mmHg) on static immersion of a subject in water. In order to counter this effect, circulatory exercises might be performed as a warm-up to the exercise session.
- The type of exercises that could be included in the main part of the programme might include breathing out into the water (EMT); resistance type exercises utilizing water resistance, flotation devices as well as weights and dynamic upper limb and trunk exercises. An IMT device might be supplied to the patient for home-based training.
- Structure of programme: 6 – 12 weeks duration (2 – 3 times per week) for a minimum of 30 minutes per session.

Recommended outcome measures to assess the effectiveness of such a programme (measured prior to and after completion):

- Ejection fraction - increase might indicate a positive outcome.
- Mean HR (% HR_{peak}) – decrease might indicate a positive outcome.
- Percentage of FEV\textsubscript{1} as well as PF – increase might indicate a positive outcome.
- PaCO\textsubscript{2} – decrease in levels might indicate a positive outcome.
- Score on 0 – 10 Borg scale – decrease in score might indicate a positive outcome.
- Exercise endurance as measured with ISWT – increase in distance walked might indicate a positive outcome.
- Quality of life measured with either the SGRQ or CRQ – increase in scores might indicate a positive outcome.

In summary, although there seems to be some evidence that hydrotherapy may improve some aspects of cardiopulmonary function in patients with COPD, increase their QoL and ability to perform ADL to some degree and contribute to psychological well-being with social integration, there are still many questions to be answered. Therefore the third hypothesis for this systematic review is rejected. One of the objectives of this review was to determine whether there was documented evidence relating to the anti-inflammatory effects of water-based exercise in patients with COPD and no evidence was found in this regard. Therefore the second hypothesis for this systematic review is rejected. Only nine articles fulfilled the inclusion criteria for this review. Most of the included articles were of limited methodological quality and therefore hypothesis one for this review is rejected.
CONCLUSIONS AND RECOMMENDATIONS

Evidence from small studies of low to moderate quality suggest that hydrotherapy may have some value in improving cardiopulmonary function, improving QoL and improving the performance of ADL in patients with COPD. The fact that hydrotherapy can offer a cost-effective group therapy alternative to land-based PR programmes for patients with COPD can perhaps be taken into consideration when community-based or hospital-based PR programmes are being implemented. Hydrotherapy programmes could be used as a primary care measure in patients with mild to moderate COPD or even in later stages of the disease when the patients are suffering from moderate to severe COPD. The introduction of an aquatic therapy programme in the early stages of COPD, perhaps in combination with patient education about the course of the disease, may lead to these patients managing their disease more effectively and perhaps slowing its progression in severity. An increase in compliance in a hydrotherapy exercise programme, as opposed to a similar land-based exercise regimen, may occur because of the variety of exercise and lack of boredom offered by the aquatic environment. The social interaction and psychological well-being derived from hydrotherapy also appears to persuade patients to comply more readily with their PR programmes. As it appears that EMT in the aquatic environment is effective in improving cardiopulmonary parameters in patients with COPD and that, the systemic muscular weaknesses and decrease in endurance levels, manifestated by the disease, can also be addressed in hydrotherapy programmes, it is perhaps feasible to consider that a combined PR programme incorporating both breathing exercises and a physical training component may be of benefit to the well-being of patients with COPD.

A significant limitation to the quality of this review is the fact that so few studies have looked at the various aspects of hydrotherapy management in patients with COPD. Those studies in existence are small and of low to moderate methodological quality. This points to an area of research which is yet to be adequately addressed. There is an urgent need for a number of well-designed randomised controlled trials to investigate the role of breathing exercises in combination with hydrotherapy physical exercise programmes on pulmonary function and exercise endurance in patients with COPD.

Investigation into the possible anti-inflammatory effects of water-based exercise in patients with COPD may also support the more frequent utilisation of this form of therapy in the future.
REFERENCES


Sarir H, Mortaz E, Janse WT, Givi ME, Nijkamp FP and Folkerts G. 2009, IL-8 production by macrophages is synergistically enhanced when cigarette smoke is combined with TNF-alpha. *Biochem Pharmacol*, Epub ahead of print. Abstract only.


APPENDIX A

APPENDIX B

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

R14/99 Ms Dorothy A Shead

CLEARANCE CERTIFICATE

PROJECT

Hydrotherapy in the Management of Chronic Diseases caused by Smoking and/or a Sedentary Lifestyle: A Systemic Review

INVESTIGATORS

Ms Dorothy A Shead.

DEPARTMENT

Physiotherapy Department

DATE CONSIDERED

2009/10/02

DECISION OF THE COMMITTEE

Unless otherwise specified, this ethical clearance is valid for 5 years and may be renewed upon application.

DATE

12/10/09

CHAIRPERSON

(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable

cc: Supervisor: Dr H van Aswegen

DECLARATION OF INVESTIGATORS

To be completed in duplicate and ONE COPY returned to the Secretary at Room 10004, 10th Floor, Senate House, University.

I/We fully understand the conditions under which I/We are authorized to carry out the above-named research and I/We guarantee to ensure compliance with these conditions. Should any departure be contemplated from the research procedure as approved I/We undertake to rethink the protocol to the Committee. I/We agree to a completion of a yearly progress report.

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...
Ms DA Shead
Postnet Suite 363
Private Bag X1
Melrose Arch
2076
South Africa

Dear Ms Shead

Master of Science in Physiotherapy: Change of title of research

I am pleased to inform you that the following change in the title of your Dissertation for the degree of has been approved:

From: Hydrotherapy in the management of chronic disease caused by smoking and/or a sedentary lifestyle: A systematic review

To: Hydrotherapy in the management of chronic obstructive pulmonary disease: A qualitative systematic review

Yours sincerely,

[Signature]

Mrs Sandra Benn
Faculty Registrar
Faculty of Health Sciences
### APPENDIX D

**LOW critical appraisal scores (DS and HvA)**

**Box 1: LOW* critical appraisal tool.**

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Can't tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Did the study address a clearly focused issue?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the population studied</td>
<td>✔️</td>
<td></td>
<td></td>
</tr>
<tr>
<td>the intervention / outcome studied</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>whether the study tried to detect a beneficial or harmful effect</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**2. Were the participants recruited in an acceptable way?**

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Can't tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were the participants representative of a defined population?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**3. Was there a sufficient number of participants selected?**

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Can't tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was there a power calculation?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did the authors provide any justification for the sample size?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**4. Was there a separate control group?**

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes (go to 5)</th>
<th>No (go to 6)</th>
<th>Can't tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>separate control group:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was there equal chance of participants being allocated to either group?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were the control representative of the intervention group (similar age, gender and variables other than the variable of interest)?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was there a sufficient number of controls selected?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Go to 7**

**5. Baseline measures for participants acting as their own controls:**

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Can't tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Where appropriate:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period?)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was the order of interventions randomised?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was the washout period between intervention/control acceptable?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**6. Were the outcomes measured accurately to minimise bias?**

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Can't tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are there references to support the use of outcome measures? (details, reliability and validity of measures)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were the measurement methods similar / the same in participants and controls?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**8. Have the confounding factors been accounted for?**

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Can't tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do the authors state potential confounding variables?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do the authors discuss and refute the impact of potential confounding variables?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have the authors taken account of the potential confounding factors in the design, results and / or in the analysis?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**9. Results**

**a. Were the results presented so the effect size was shown or could be calculated?**

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Can't tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are means / SD (or the raw-data) available to allow calculation of effect size?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size of the p value</td>
<td>✔️</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size of the confidence intervals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are data for participant attrition / withdrawal presented?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**b. Do you subjectively believe the results?**

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Can't tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOT do you accept the results?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What are the bottom line results?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was the analysis appropriate to the design?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safety with beneficial outcomes</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Final score**

*LOW = Lewis, Olid, Williams. Yes = 1, No = 0, Can't tell = 0

---

*Appendix D continued...*
**Box 1: LOW* critical appraisal tool.**

1. Did the study address a clearly focused issue?  
   - the population studied  
   - the intervention / outcome studied  
   - whether the study tried to detect a beneficial or harmful effect  
   - Yes  
   - No  
   - Can't tell

2. Were the participants recruited in an acceptable way?  
   - Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?  
   - Were the participants representative of a defined population?  
   - How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)  
   - No exclusion criteria  
   - Yes  
   - No  
   - Can't tell

3. Was there a sufficient number of participants selected?  
   - Was there a power calculation?  
   - Did the authors provide any justification for the sample size?  
   - Yes  
   - No  
   - Can't tell

4. Was there a separate control group?  
   - Yes (go to 5)  
   - No (go to 6)

   **Separate control group:**  
   - Was there equal chance of participants being allocated to either group?  
   - Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?  
   - Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?  
   - Was there a sufficient number of controls selected?  
   - Yes  
   - No  
   - Can't tell

5. Baseline measures for participants acting as their own controls:  
   - Where appropriate:  
   - Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period)?  
   - Was the order of interventions randomised?  
   - Was the washout period between intervention/control acceptable?  
   - Yes  
   - No  
   - Can't tell

6. Were the outcomes measured accurately to minimise bias?  
   - Are there references to support the use of outcome measures? (details, reliability and validity of measures)  
   - Were the measurement methods similar / the same in participants and controls?  
   - Yes  
   - No  
   - Can't tell

8. Have the confounding factors been accounted for?  
   - Do the authors state potential confounding variables?  
   - Do the authors discuss and refute the impact of potential confounding variables?  
   - Have the authors taken account of the potential confounding factors in the design, results and / or in the analyses?  
   - Yes  
   - No  
   - Can't tell

9. Results  
   a. Were the results presented so the effect size was shown or could be calculated?  
      - Mean / SD (or the raw data) available to allow calculation of effect size  
      - Size of the p value  
      - Size of the confidence intervals  
      - Are data for participant attrition / withdrawal presented?  
      - Yes  
      - No  
      - Can't tell
   b. Do you subjectively believe the results?  
      - NOT do you accept the results?  
      - What are the bottom line results?  
      - Was the analysis appropriate to the design?  
      - Yes  
      - No  
      - Can't tell

**Final score**  
2 / 9

*LOW = Lewis, Olds, Williams. Yes = 1, No = 0, Can't tell = 0*
Box 1: LOW* critical appraisal tool.

1. Did the study address a clearly focused issue?
   - the population studied [X]
   - the intervention / outcome studied [X]
   - whether the study tried to detect a beneficial or harmful effect [X]
     Yes  No  Can't tell

2. Were the participants recruited in an acceptable way?
   - Were the eligibility criteria clearly specified so that the participant recruitment could be repeated? [X]
   - Were the participants representative of a defined population? [X]
     Yes  No  Can't tell
   - How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention) [X]
     Bias was likely

3. Was there a sufficient number of participants selected?
   - Was there a power calculation? [No]
   - Did the authors provide any justification for the sample size? [No]
     Yes  No  Can't tell

4. Was there a separate control group? [No]
   Yes (go to 5)
   No (go to 6)

   - separate control group:
     - Was there equal chance of participants being allocated to either group? [Yes]
     - Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)? [No]
     - Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated? [No]
     - Was there a sufficient number of controls selected? [No]
     Go to 7

5. Baseline measures for participants acting as their own controls:
   Where appropriate:
   - Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period?) for SO2, NO2, NO, for blood gas, COX and pulmonary vasculopathy? [No]
   - Was the order of interventions randomised? [No]
   - Was the washout period between interventions/controls acceptable? [N/A]
   Yes  No  Can't tell

6. Were the outcomes measured accurately to minimise bias?
   - Are there references to support the use of outcome measures? (details, reliability and validity of measures) [N/A]
     Yes  No  Can't tell
   - Were the measurement methods similar / the same in participants and controls? [N/A]

7. Have the confounding factors been accounted for?
   - Do the authors state potential confounding variables? [No]
   - Do the authors discuss and refute the impact of potential confounding variables? [No]
     Yes  No  Can't tell
   - Have the authors taken account of the potential confounding factors in the design, results and / or in the analyses? [No]

9. Results
a. Were the results presented so the effect size was shown or could be calculated?
   - Are mean / SD (or the raw data) available to allow calculation of effect size [Yes]
   - Size of the p value [Given for resp, abg, EF, LV, AUC results]
   - Size of the confidence intervals [No]
     Yes  No  Can't tell
   - Are data for participant attrition / withdrawal presented? [No]

b. Do you subjectively believe the results?
   - NOT do you accept the results? [No]
   - What are the bottom line results? [Yes]
     Data for FVC (%), FEV1% replicated from 1997
     8% O2 vs. 6% CO2, replicated from 1997 n=22
   - Was the analysis appropriate to the design? [Yes]
     Can't tell

Final score: 2 / 5

*LOW = Lewis, Olds, Williams. Yes = 1, No = 0, Can't tell = 0
<table>
<thead>
<tr>
<th>Box 1: LOW® critical appraisal tool:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Did the study address a clearly focused issue?</td>
</tr>
<tr>
<td>• the population studied  ✔</td>
</tr>
<tr>
<td>• the intervention / outcome studied  ✔</td>
</tr>
<tr>
<td>• whether the study tried to detect a beneficial or harmful effect</td>
</tr>
<tr>
<td>[beneficial related to different lengths of intervention]</td>
</tr>
<tr>
<td>Yes  ✔</td>
</tr>
<tr>
<td>No  ◐</td>
</tr>
<tr>
<td>Can't tell</td>
</tr>
</tbody>
</table>

2. Were the participants recruited in an acceptable way? |
• Were the eligibility criteria clearly specified so that the participant recruitment could be repeated? [No exclusion criteria step 1 (and 2)]
• Were the participants representative of a defined population?  ✔
• How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)  ◐
• Was there a sufficient number of participants selected?  No  ◐
• Was the authors provide any justification for the sample size?  No  ◐

3. Was there a separate control group?  No  ◐

4. Was there a separate control group?  No  ◐
   (No, go to 5)

5. Separate control group:
   • Was there equal chance of participants being allocated to either group?  No  ◐
   • Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?  No  ◐
   • Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?  No  ◐
   • Was there a sufficient number of controls selected?  No  ◐
   Go to 7

6. Baseline measures for participants acting as their own controls:
   Where appropriate:
   • Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period)?  No  ◐
   • Was the order of interventions randomised?  No  ◐
   • Was the washout period between intervention/control acceptable?  No  ◐

7. Were the outcomes measured accurately to minimise bias?
   Are there references to support the use of outcome measures? (details, reliability and validity of measures)  ✔
   • Were the measurement methods similar / the same in participants and controls?  ✔

8. Have the confounding factors been accounted for?
   • Do the authors state potential confounding variables?  No  ◐
   • Do the authors discuss and refute the impact of potential confounding variables?  No  ◐
   • Have the authors taken account of the potential confounding factors in the design, results and/or in the analyses?  No  ◐

9. Results
   a. Were the results presented so the effect size was shown or could be calculated?
      • Are mean / SD (or the raw data) available to allow calculation of effect size  ✔
      • Size of the p value  ✔
      • Size of the confidence intervals  ✔
      • Are data for participant attrition / withdrawal presented?  ✔
   b. Do you subjectively believe the results?
      • NOT do you accept the results?  ✔
      • What are the bottom line results?  ✔
      • Was the analytic appropriate to the design?  ✔
   Final score 3 / 9

*LOW = Lewis, Olds, Williams. Yes = 1, No = 0, Can't tell = 0
Box 1: LOW* critical appraisal tool.

1. Did the study address a clearly focused issue?
   - the population studied
   - the intervention / outcome studied
   - whether the study tried to detect a beneficial or harmful effect
   Yes
   No
   Can't tell

2. Were the participants recruited in an acceptable way?
   - Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?
   - Were the participants representative of a defined population?
   - How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)
   Yes
   No
   Can't tell

3. Was there a sufficient number of participants selected?
   - Was there a power calculation?
   - Did the authors provide any justification for the sample size?
   Yes
   No
   Can't tell

4. Was there a separate control group?
   No
   Yes (go to 5)
   (No (go to 6))

   separate control group:
   - Was there equal chance of participants being allocated to either group?
   - Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?
   Yes
   No
   Can't tell

5. Baseline measures for participants acting as their own controls:
   Where appropriate:
   - Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period)?
   - Was the order of interventions randomised?
   Yes
   No
   Can't tell

6. Were the outcomes measured accurately to minimise bias?
   - Were there references to support the use of outcome measures? (details, reliability and validity of measures)
   - Were the measurement methods similar / the same in participants and controls?
   Yes
   No
   Can't tell

7. Have the confounding factors been accounted for?
   - Do the authors state potential confounding variables?
   - Do the authors discuss and refute the impact of potential confounding variables?
   - Have the authors taken account of the potential confounding factors in the design, results and / or in the analyses?
   Yes
   No
   Can't tell

8. Results
   a. Were the results presented so the effect size was shown or could be calculated?
      - Are mean / SD (or the raw data) available to allow calculation of effect size
      - Size of the p value
      - Size of the confidence intervals
      - Are data for participant attrition / withdrawal presented?
      Yes
      No
      Can't tell

   b. Do you subjectively believe the results?
      - NCT do you accept the results?
      - What are the bottom line results?
      - Was the analysis appropriate to the design?
      Yes
      No
      Can't tell

Final score

*LOW = Lewis, Olds, Williams. Yes = 1, No = 0, Can't tell = 0
### Box 1: LOW* critical appraisal tool.

1. **Did the study address a clearly focused issue?**
   - the population studied
   - the intervention / outcome studied
   - whether the study tried to detect a beneficial or harmful effect

2. **Were the participants recruited in an acceptable way?**
   - Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?
   - Were the participants representative of a defined population?
   - How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)

3. **Was there a sufficient number of participants selected?**
   - Was there a power calculation?
   - Did the authors provide any justification for the sample size?

4. **Was there a separate control group?**
   - Separate control group:
     - Was there equal chance of participants being allocated to either group?
     - Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?
     - Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?
     - Was there a sufficient number of controls selected?

5. **Baseline measures for participants acting as their own controls:**
   - Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period)?
   - Was the order of interventions randomised?
   - Was the washout period between intervention/control acceptable?

6. **Were the outcomes measured accurately to minimise bias?**
   - Are there references to support the use of outcome measures? (details, reliability and validity of measures)
   - Were the measurement methods similar / the same in participants and controls?

7. **Have the confounding factors been accounted for?**
   - Do the authors state potential confounding variables?
   - Do the authors discuss and refute the impact of potential confounding variables?
   - Have the authors taken account of the potential confounding factors in the design, results and / or in the analyses?

8. **Results**
   - Were the results presented so the effect size was shown or could be calculated?
     - Are mean / SD (or the raw data) available to allow calculation of effect size
     - Size of the p value
     - Size of the confidence interval
     - Are data for participant attrition / withdrawal presented?
   - Do you subjectively believe the results?
     - NOT do you accept the results?
     - What are the bottom line results? (high intensity physical group having 20%)
     - Was the analysis appropriate to the design?

---

*LOW = Lewis, Olde, Williams. Yes = 1, No = 0, Can't tell = 0

**Final score:** 8 / 9
### Box 1: LOW* critical appraisal tool

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Can't tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Did the study address a clearly focussed issue?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• the population studied</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• the intervention / outcome studied</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• whether the study tried to detect a beneficial or harmful effect</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Were the participants recruited in an acceptable way?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Were the participants representative of a defined population?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)</td>
<td>Can't tell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Was there a sufficient number of participants selected?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Was there a power calculation?</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Did the authors provide any justification for the sample size?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Was there a separate control group?</td>
<td>Yes (go to 5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• separate control group:</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?</td>
<td>Can't tell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Was there a sufficient number of controls selected?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Baseline measures for participants acting as their own controls:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Where appropriate:</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period)?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Was the order of interventions randomised?</td>
<td>Can't tell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Was the washout period between intervention/control acceptable?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Were the outcomes measured accurately to minimise bias?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Are there references to support the use of outcome measures? (details, reliability and validity of measures)</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Were the measurement methods similar / the same in participants and controls?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Have the confounding factors been accounted for?</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Do the authors state potential confounding variables?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Do the authors discuss and refute the impact of potential confounding variables?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Have the authors taken account of the potential confounding factors in the design, results and / or in the analyses?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Results</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Were the results presented so the effect size was shown or could be calculated?</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Are mean / SD (or the raw data) available to allow calculation of effect size?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Size of the p value</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Size of the confidence intervals</td>
<td>Can't tell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Are data for participant inclusion / withdrawal presented</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Do you subjectively believe the results?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• NOT do you accept the results?</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• What are the bottom line results?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Was the analysis appropriate to the design?</td>
<td>Can't tell</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*LOW = Lewis, Olds, Williams. Yes = 1, No = 0, Can't tell = 0
Box 1: LOW critical appraisal tool

1. Did the study address a clearly focused issue?
   - the population studied
   - the intervention / outcome studied
   - whether the study tried to detect a beneficial or harmful effect?
   Yes
   No
   Can't tell

2. Were the participants recruited in an acceptable way?
   -Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?
   Yes
   No
   Can't tell
   -Were the participants representative of a defined population?
   Yes
   No
   Can't tell
   -How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)
   Bias likely

3. Was there a sufficient number of participants selected?
   -Was there a power calculation?
     No
     Yes
     Can't tell
   -Did the authors provide any justification for the sample size?
     No
     Yes
     Can't tell

4. Was there a separate control group?
   Yes (go to 5)
   No (go to 6)

   _Separate control group:_
   -Was there equal chance of participants being allocated to either group?
   Yes
   No
   Can't tell
   -Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?
   Yes
   No
   Can't tell
   -Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?
   Yes
   No
   Can't tell
   -Was there a sufficient number of controls selected?
   Yes
   No
   Can't tell

5. Baseline measures for participants acting as their own controls:
   -Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period)?
   Yes
   No
   Can't tell
   -Was the order of interventions randomised?
   Yes
   No
   Can’t tell
   -Was the washout period between intervention/control acceptable?
   N/A

6. Were the outcomes measured accurately to minimise bias?
   -Are there references to support the use of outcome measures? (details, reliability and validity of measures)
     Yes
     No
     Can't tell
   -Were the measurement methods similar / the same in participants and controls?
     Yes
     No
     Can't tell

7. Have the confounding factors been accounted for?
   -Do the authors state potential confounding variables?
     Yes
     No
     Can’t tell
   -Do the authors discuss and refute the impact of potential confounding variables?
     Yes
     No
     Can’t tell
   -Have the authors taken account of the potential confounding factors in the design, results and / or in the analyses?
     Yes
     No
     Can’t tell

8. Results
   a. Were the results presented so the effect size was shown or could be calculated?
      -Are mean / SD (or the raw-data) available to allow calculation of effect size?
        Yes
        No
        Can't tell
   b. Do you subjectively believe the results?
      -NOT do you accept the results?
        Yes
        No
        Can’t tell
      -What are the bottom line results? MT + hydrotherapy more effective than PRO
        Yes
        No
        Can’t tell

   Final score

*LOW = Lewis, Oles, Williams. Yes = 1, No = 0, Can't tell = 0
Box 1: LOW* critical appraisal tool.

1. Did the study address a clearly focused issue?  
   - the population studied  ✔  
   - the intervention / outcome studied  ✔  
   - whether the study tried to detect a beneficial or harmful effect  ✔ Gone.
   - Yes  
   - No  
   - Can't tell

2. Were the participants recruited in an acceptable way?  
   - Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?  
   - Were the participants representative of a defined population?  
   - How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)  
   - Yes  
   - No  
   - Can't tell

3. Was there a sufficient number of participants selected?  
   - Was there a power calculation?  
   - Did the authors provide any justification for the sample size?  
   - Yes  
   - (No)  
   - Can't tell

4. Was there a separate control group?  
   - Yes (go to 5)  
   - No (go to 6)

   - Separate control group:  
     - Was there equal chance of participants being allocated to either group?  
     - Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?  
     - Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?  
     - Yes  
     - No  
     - Can't tell

5. Baseline measures for participants acting as their own controls:  
   - Were the baseline stable (how confident are you that the pre measures were stable, was there a run-in period)?  
   - Was the order of interventions randomised?  
   - Was the washout period between intervention/control acceptable?  
   - Yes  
   - No  
   - Can't tell

6. Were the outcomes measured accurately to minimise bias?  
   - Are these references to support the use of outcome measures? (details, reliability and validity of measures)  
   - Yes  
   - No  
   - Can't tell

7. Were the measurement methods similar in participants and controls?  
   - Yes  
   - No  
   - Can't tell

8. Have the confounding factors been accounted for?  
   - Do the authors state potential confounding variables?  
   - Do the authors discuss and refute the impact of potential confounding variables?  
   - Have the authors taken account of the potential confounding factors in the design, results and / or in the analysis?  
   - Yes  
   - (No)  
   - Can't tell

9. Results  
   a. Were the results presented so the effect size was shown or could be calculated?  
   - Yes  
   - No  
   - Can't tell

   b. Do you subjectively believe the results?  
   - Yes  
   - No  
   - Can't tell

Final score  

*LOW = Lewis, Olis, Williams. Yes = 1, No = 0, Can't tell = 0
### Box 1: LOW* critical appraisal tool.

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Can't tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did the study address a clearly focused issue?</td>
<td>1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• the population studied</td>
<td>✔</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• the intervention / outcome studied</td>
<td>✔</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• whether the study tried to detect a beneficial or harmful effect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were the participants recruited in an acceptable way?</td>
<td>✔</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?</td>
<td>✔</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Were the participants representative of a defined population?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was there a sufficient number of participants selected?</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>• Was there a power calculation? X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Did the authors provide any justification for the sample size?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was there a separate control group?</td>
<td>Yes (go to 5)</td>
<td>No (go to 6)</td>
<td></td>
</tr>
<tr>
<td>Separate control group:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Was there equal chance of participants being allocated to either group?</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>• Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Was there a sufficient number of controls selected?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline measures for participants acting as their own controls:</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>Where appropriate:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Was there a baseline stable (how confident are you that the pre measures were stable, was there a run-in period)?</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>• Was the order of interventions randomised? X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Was the washout period between intervention/control acceptable?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The outcomes measured accurately to minimise bias?</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>• Are there references to support the use of outcome measures? (details, reliability and validity of measures)</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>• Were the measurement methods similar / the same in participants and controls?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have the confounding factors been accounted for?</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>• Do the authors state potential confounding variables? X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Do the authors discuss and refute the impact of potential confounding variables? X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Have the authors taken account of the potential confounding factors in the design, results and / or in the analysis? X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>a. Were the results presented so the effect size was shown or could be calculated?</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>• Are mean / SD (or the raw data) available to allow calculation of effect size</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>• Size of the p value</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>• Size of the confidence interval X</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>• Are data for participant attrition/withdrawal presented?</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>• Do you subjectively believe the results?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• What are the bottom line results?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Was the analysis appropriate to the design?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final score</td>
<td>5 / 9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Box 1: LOW* critical appraisal tool.**

1. Did the study address a clearly focused issue?
   - the population studied
   - the intervention / outcome studied
   - whether the study tried to detect a beneficial or harmful effect
   **(Yes)  No  Can’t tell**

2. Were the participants recruited in an acceptable way?
   - Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?
   - Were the participants representative of a defined population?
   - How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)
   **(Yes)  (No)  Can’t tell**

3. Was there a sufficient number of participants selected?
   - Was there a power calculation?
   - Did the authors provide any justification for the sample size?
   **(Yes)  (No)  Can’t tell**

4. Was there a separate control group?
   **(Yes) (go to 5)  (No) (go to 6)**

5. Separate control group:
   - Was there equal chance of participants being allocated to either group?
   - Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?
   - Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?
   - Was there a sufficient number of controls selected?
   **(Yes)  (No)  Can’t tell**

6. Baseline measures for participants acting as their own controls:
   - Where appropriate:
     - Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period?)
     - Was the order of interventions randomised?
     - Was the washout period between intervention/control acceptable?
   **(Yes)  (No)  Can’t tell**

7. Were the outcomes measured accurately to minimise bias?
   - Are there references to support the use of outcome measures? (details, reliability and validity of measure)
   - Were the measurement methods similar / the same in participants and controls?
   **(Yes)  (No)  Can’t tell**

8. Have the confounding factors been accounted for?
   - Do the authors state potential confounding variables?
   - Do the authors discuss and refute the impact of potential confounding variables?
   - Have the authors taken account of the potential confounding factors in the design, results and/or in the analyses?
   **(Yes)  (No)  Can’t tell**

9. Results
   a. Were the results presented so the effect size was shown or could be calculated?
     - Are mean / SD (or the raw data) available to allow calculation of effect size?
     - Size of the p value
     - Size of the confidence intervals
   - Are data for participant attrition/withdrawal presented?
   **(Yes)  (No)  Can’t tell**

   b. Do you subjectively believe the results?
   - NOT do you accept the results?
   - What are the bottom line results?
   - Was the analysis appropriate to the design?
   **(Yes)  (No)  Can’t tell**

   **Final score: 2/9**

*LOW = Lewis, Ols, Williams. Yes = 1, No = 0, Can’t tell = 0*
Box 1: LOW* critical appraisal tool.

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Can't tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Did the study address a clearly focused issue?</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- the population studied</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- the intervention/outcome studied</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- whether the study tried to detect a beneficial or harmful effect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Were the participants recruited in an acceptable way?</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Were the participants representative of a defined population?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Was there a sufficient number of participants selected?</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- Was there a power calculation? X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Did the authors provide any justification for the sample size? X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Was there a separate control group?</td>
<td>Yes</td>
<td>No (go to 5)</td>
<td></td>
</tr>
<tr>
<td>- Separate control group:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Were there equal chances of participants being allocated to either group?</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Was there a sufficient number of controls selected?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Go to 7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Baseline measures for participants acting as their own controls:</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>Where appropriate:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period)?</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- Was the order of interventions randomised?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Was the washout period between intervention/control acceptable?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Were the outcomes measured accurately to minimise bias?</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- Are there references to support the use of outcome measures? (details, reliability and validity of measures)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Were the measurement methods similar / the same in participants and controls?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Have the confounding factors been accounted for?</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- Do the authors state potential confounding variables? X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Do the authors discuss and refute the impact of potential confounding variables? X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Have the authors taken account of the potential confounding factors in the design, results and / or in the analyses? X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Results</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>a. Were the results presented so the effect size was shown or calculated?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Are mean / SD (or the raw data) available to allow calculation of effect size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Size of the p value</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Size of the confidence intervals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Are data for participant attrition/withdrawal presented? X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you subjectively believe the results?</td>
<td>Yes</td>
<td>No</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- NOT do you accept the results?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- What are the bottom line results?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Was the analysis appropriate to the design?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Box 1: LOW* critical appraisal tool.

1. Did the study address a clearly focused issue?
   - the population studied
   - the intervention / outcome studied
   - whether the study tried to detect a beneficial or harmful effect
     - Yes
     - No
     - Can't tell

2. Were the participants recruited in an acceptable way?
   - Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?
     - Yes
     - No
     - Can't tell
   - Were the participants representative of a defined population?
     - Yes
     - No
     - Can't tell
   - How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)
     - Likely

3. Was there a sufficient number of participants selected?
   - Was there a power calculation? x
     - Yes
     - No
     - Can't tell
   - Did the authors provide any justification for the sample size? x

4. Was there a separate control group?
   - Yes (go to 6)
   - No (go to 7)

5. Separate control group:
   - Were equal chances of participants being allocated to either group?
     - Yes
     - No
     - Can't tell
   - Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?
     - Yes
     - No
     - Can't tell
   - Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?
     - Yes
     - No
     - Can't tell
   - Was there a sufficient number of controls selected?
     - Yes
     - No
     - Can't tell

6. Baseline measures for participants acting as their own controls:
   - Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period)?
     - Yes
     - No
     - Can't tell
   - Was the order of interventions randomised? Can't tell how interventions were done.
   - Was the washout period between intervention/control acceptable? Can't tell

7. Were the outcomes measured accurately to minimise bias?
   - Are there references to support the use of outcome measures? (details, reliability and validity of measures)
     - Yes
     - No
     - Can't tell
   - Were the measurement methods similar / the same in participants and controls?
     - Yes
     - No
     - Can't tell

8. Have the confounding factors been accounted for?
   - Do the authors state potential confounding variables? x
     - Yes
     - No
     - Can't tell
   - Do the authors discuss and refute the impact of potential confounding variables? x
     - Yes
     - No
     - Can't tell
   - Have the authors taken account of the potential confounding factors in the design, results and / or in the analysis?
     - Yes
     - No
     - Can't tell

9. Results
   a. Were the results presented so the effect size was shown or could be calculated?
     - Yes
     - No
     - Can't tell
   b. Are mean / SD (or the raw data) available to allow calculation of effect size?
     - Yes
     - No
     - Can't tell
   c. Size of the p value
     - Yes
     - No
     - Can't tell
   d. Size of the confidence intervals
     - Yes
     - No
     - Can't tell
   e. Are data for participant attrition / withdrawal presented?
     - Yes
     - No
     - Can't tell
   f. Do you subjectively believe the results?
     - Yes
     - No
     - Can't tell
   g. NOT do you accept the results?
     - Yes
     - No
     - Can't tell
   h. What were the bottom line results?
     - Yes
     - No
     - Can't tell
   i. Was the analysis appropriate to the design?
     - Yes
     - No
     - Can't tell

Final score 9 / 10

*LOW = Lewis, Odds, Williams. Yes = 1, No = 0, Can't tell = 0
### Box 1. LOW* critical appraisal tool.

1. Did the study address a clearly focused issue?
   - the population studied?
   - the intervention / outcome studied?
   - whether the study tried to detect a beneficial or harmful effect?
   - Yes
   - No
   - Can’t tell

2. Were the participants recruited in an acceptable way?
   - Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?
   - Were the participants representative of a defined population?
   - How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)?
   - Yes
   - No
   - Can’t tell

3. Was there a sufficient number of participants selected?
   - Was there a power calculation?
   - Did the authors provide any justification for the sample size?
   - Yes
   - No
   - Can’t tell

4. Was there a separate control group?
   - Yes (go to 5)
   - No (go to 6)

5. Separate control group:
   - Was there equal chance of participants being allocated to either group?
   - Were the controls representative of the intervention group (similar age, gender, and variables other than the variable of interest)?
   - Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?
   - Was there a sufficient number of controls selected?
   - Yes
   - No
   - Can’t tell

6. Baseline measures for participants acting as their own controls:
   Where appropriate:
   - Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period)?
   - Was the order of interventions randomized?
   - Was the washout period between intervention/control acceptable?
   - Yes
   - No
   - Can’t tell

7. Were the outcomes measured accurately to minimise bias?
   - Are there references to support the use of outcome measures (details, reliability and validity of measures)?
   - Were the measurement methods similar / the same in participants and controls?
   - Yes
   - No
   - Can’t tell

8. Have the confounding factors been accounted for?
   - Do the authors state potential confounding variables?
   - Do the authors discuss and refute the impact of potential confounding variables?
   - Have the authors taken account of the potential confounding factors in the design, results and / or in the analyses?
   - Yes
   - No
   - Can’t tell

9. Results
   a. Were the results presented so the effect size was shown or could be calculated?
   - Are mean / SD (or the raw data) available to allow calculation of effect size?
   - Size of the p value?
   - Size of the confidence intervals?
   - Are data for participant attrition / withdrawal presented?
   - Yes
   - No
   - Can’t tell

   b. Do you subjectively believe the results?
   - NOT do you accept the results?
   - What are the bottom line results?
   - Was the analysis appropriate to the design?
   - Can’t tell

Final score 5 / 9

*LOW = Low, Olde, Williams. Yes = 1, No = 0, Can’t tell = 0
Box 1: LOW* critical appraisal tool.

1. Did the study address a clearly focused issue?
   - the population studied
   - the intervention/outcome studied
   - whether the study tried to detect a beneficial or harmful effect
     - Yes
     - No
     - Can't tell

2. Were the participants recruited in an acceptable way?
   - Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?
   - Were the participants representative of a defined population?
   - How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)
     - Yes
     - No
     - Can't tell

3. Was there a sufficient number of participants selected?
   - Was there a power calculation?
   - Did the authors provide any justification for the sample size?
     - Yes
     - No
     - Can't tell

4. Was there a separate control group?
   - Yes (go to 5)
   - No (go to 5)

5. Separate control group:
   - Was there equal chance of participants being allocated to either group?
   - Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?
   - Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?
   - Were there a sufficient number of controls selected?
     - Yes
     - No
     - Can't tell

6. Baseline measures for participants acting as their own controls:
   - Where appropriate:
     - Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period?)
     - Was the order of interventions randomised?
     - Was the washout period between intervention/control acceptable?
     - Yes
     - No
     - Can't tell

7. Were the outcomes measured accurately to minimise bias?
   - Are there references to support the use of outcome measures? (details, reliability and validity of measures)
   - Were the measurement methods similar/the same in participants and controls?
     - Yes
     - No
     - Can't tell

8. Have the confounding factors been accounted for?
   - Do the authors state potential confounding variables?
   - Do the authors discuss and refute the impact of potential confounding variables?
   - Have the authors taken account of the potential confounding factors in the design, results and/or in the analyses?
     - Yes
     - No
     - Can't tell

9. Results
   - Were the results presented so the effect size was shown or could be calculated?
     - Yes
     - No
     - Can't tell
   - Are mean/SD (or the raw data) available to allow calculation of effect size?
   - Size of the p value
   - Size of the confidence intervals (CI or ES estimates)
   - Are data for participant attrition/withdrawal presented?
   - Do you subjectively believe the results?
     - Yes
     - No
   - What are the bottom line results?
     - Can't tell
   - Was the analysis appropriate to the design?
     - Can't tell

Final score

*LOW = Lewis, O'Shea, Williams. Yes = 1, No = 0, Can't tell = 0
Box 1: LOW* critical appraisal tool.

1. Did the study address a clearly focused issue?
   - the population studied
   - the intervention / outcome studied
   - whether the study tried to detect a beneficial or harmful effect

2. Were the participants recruited in an acceptable way?
   - Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?
   - Were the participants representative of a defined population?
   - How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)

3. Was there a sufficient number of participants selected?
   - Was there a power calculation?
   - Did the authors provide any justification for the sample size?

4. Was there a separate control group?
   - Yes (go to 5)
   - No (go to 6)

5. Separate control group:
   - Were there equal chance of participants being allocated to either group? 
   - Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?
   - Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?
   - Was there a sufficient number of controls selected?

6. Baseline measures for participants acting as their own controls:
   - Were the baseline measures stable (how confident are you that the pre measures were stable, was there a run-in period?)
   - Was the order of interventions randomised?
   - Was the washout period between intervention/control acceptable?

7. Were the outcomes measured accurately to minimise bias?
   - Are there references to support the use of outcome measures? (details, reliability and validity of measures)
   - Were the measurement methods similar / the same in participants and controls?

8. Have the confounding factors been accounted for?
   - Do the authors state potential confounding variables?
   - Do the authors discuss and refute the impact of potential confounding variables?
   - Have the authors taken account of the potential confounding factors in the design, results and/or in the analysis?

9. Results
   a. Were the results presented so the effect size was shown or could be calculated?
   - Are mean / SD (or the raw data) available to allow calculation of effect size?
   - Size of the p value
   - Size of the confidence intervals
   - Are data for participant adherence / withdrawal presented?
   b. Do you subjectively believe the results?
   - NOT do you accept the results?
   - What are the bottom line results?
   - Was the analysis appropriate to the design?

Final score

*LOW = Lewis, Odds, Williams. Yes = 1, No = 0, Can't tell = 0
### Box 1: LOW* critical appraisal tool.

1. Did the study address a clearly focused issue?
   - the population studied
   - the intervention / outcome studied
   - whether the study tried to detect a beneficial or harmful effect

   **Yes**
   **No**
   **Can’t tell**

2. Were the participants recruited in an acceptable way?
   - Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?
   - Were the participants representative of a defined population?
   - How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)

   **Yes**
   **No**
   **Can’t tell**

3. Was there a sufficient number of participants selected?
   - Was there a power calculation?
   - Did the authors provide any justification for the sample size?

   **Yes**
   **No**
   **Can’t tell**

4. Was there a separate control group?

   **Yes (go to 5)**
   **No (go to 6)**

5. Separate control group:
   - Was there equal chance of participants being allocated to either group?
   - Were the controls representative of the intervention group (similar age, gender and variables other than the variable of interest)?
   - Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?
   - Was there a sufficient number of controls selected?

6. Baseline measures for participants acting as their own controls:
   Where appropriate:
   - Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period?)
   - Was the order of interventions randomised?

   **Yes**
   **No**
   **Can’t tell**

7. Were the outcomes measured accurately to minimise bias?
   - Are there references to support the use of outcome measures? (details, reliability and validity of measures)
   - Were the measurement methods similar / the same in participants and controls?

   **Yes**
   **No**
   **Can’t tell**

8. Have the confounding factors been accounted for?
   - Do the authors state potential confounding variables?
   - Do the authors discuss and refute the impact of potential confounding variables?
   - Have the authors taken account of the potential confounding factors in the design, results and / or in the analyses?

   **Yes**
   **No**
   **Can’t tell**

9. Results
   a. Were the results presented so the effect size was shown or could be calculated?
      - Are mean / SD (or the raw data) available to allow calculation of effect size?
      - Size of the p value
      - Size of the confidence intervals
      - Are data for participant attrition / withdrawal presented?

   **Yes**
   **No**
   **Can’t tell**

   b. Do you subjectively believe the results?
      - NOT do you accept the results?
      - What are the bottom line results?
      - Was the analysis appropriate to the design?

   **Yes**
   **No**
   **Can’t tell**

**Final score**

*LOW = Lewis, Olds, Williams. Yes = 1, No = 0, Can’t tell = 0
Box 1: LOW* critical appraisal tool.

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Can't tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Did the study address a clearly focused issue?</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- the population studied</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- the intervention / outcome studied</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- whether the study tried to detect a beneficial or harmful effect</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>2. Were the participants recruited in an acceptable way?</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- Were the eligibility criteria clearly specified so that the participant recruitment could be repeated?</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- Were the participants representative of a defined population?</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- How likely was the recruitment process to introduce bias? (participants likely to respond positively or negatively to intervention)</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>3. Was there a sufficient number of participants selected?</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- Was there a power calculation?</td>
<td>❌</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Did the authors provide any justification for the sample size?</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>4. Was there a separate control group?</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(go to 5)</td>
<td>No</td>
<td>(go to 6)</td>
<td></td>
</tr>
<tr>
<td>5. Separate control group:</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Were there equal chance of participants being allocated to either group?</td>
<td>No</td>
<td></td>
<td>Can't tell</td>
</tr>
<tr>
<td>- Were the controls representative of the intervention group?</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>(similar age, gender and variables other than the variable of interest)?</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- Were the eligibility criteria clearly specified so that the recruitment of the controls could be repeated?</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>- Was there a sufficient number of controls selected?</td>
<td>✔</td>
<td>❌</td>
<td>Can't tell</td>
</tr>
<tr>
<td>Go to 7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Baseline measures for participants acting as their own controls:</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Where appropriate:</td>
<td>No</td>
<td></td>
<td>Can’t tell</td>
</tr>
<tr>
<td>- Was the baseline stable (how confident are you that the pre measures were stable, was there a run-in period?)</td>
<td>✔</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>- Was the order of interventions randomised?</td>
<td>✔</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>- Was the washout period between intervention/control acceptable?</td>
<td>✔</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>7. Were the outcomes measured accurately to minimise bias?</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Are there references to support the use of outcome measures? (details, reliability and validity of measures)</td>
<td>No</td>
<td></td>
<td>Can’t tell</td>
</tr>
<tr>
<td>- Were the measurement methods similar / the same in participants and controls?</td>
<td>✔</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>8. Have the confounding factors been accounted for?</td>
<td>Yes</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>- Do the authors state potential confounding variables?</td>
<td>✔</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>- Do the authors discuss and refute the impact of potential confounding variables?</td>
<td>✗</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Have the authors taken account of the potential confounding factors in the design, results and / or in the analyses?</td>
<td>✔</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>9. Results</td>
<td>Yes</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>a. Were the results presented so the effect size was shown or could be calculated?</td>
<td>✔</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>- Are mean / SD (or the raw data) available to allow calculation of effect size?</td>
<td>✔</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>- Size of the p value</td>
<td>✔</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>- Size of the confidence intervals</td>
<td>✔</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>- Are data for participant attrition / withdrawal presented?</td>
<td>✔</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>b. Do you subjectively believe the results?</td>
<td>Yes</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>- NOT do you accept the results?</td>
<td>❌</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>- What are the bottom line results?</td>
<td>❌</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>- Was the analysis appropriate to the design?</td>
<td>❌</td>
<td>❌</td>
<td>Can’t tell</td>
</tr>
<tr>
<td>Fiscal score</td>
<td>7 / 9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*LOW = Lewis, Olds, Williams. Yes = 1, No = 0, Can’t tell = 0