Economic evaluation of domiciliary long-term non-invasive positive pressure ventilation in Duchenne muscular dystrophy

Stephen James Colgan

Submitted in partial fulfilment of the requirements of the degree of Doctor of Philosophy

August 2015

School of Public Health
Department of Medicine, St Vincent's Hospital Melbourne

The University of Melbourne

Produced on archival quality paper
Abstract

Duchenne muscular dystrophy (DMD) is a rare X-linked recessive, incurable and invariably fatal disease that affects approximately one in every 3300 live male births. It is the result of the failure of the body to produce a protein called dystrophin that stabilises and protects the plasma membranes of muscle cells from the mechanical stresses induced by muscle fibre contraction. In the absence of this protein muscle contraction disrupts the cell membranes, triggering a process that results in cell death and a reduction in muscle strength and function.

There is a rapid decline in respiratory muscle strength approximately 1 year after the loss of independent ambulation. The first manifestations of this are seen during sleep, which is disrupted by periods of abnormal breathing (sleep disordered breathing). As respiratory function worsens this disruption increases in frequency and severity and unless some form of long-term mechanical ventilation (LTMV) is used to support respiratory function, it inevitably results in respiratory failure and death in the early 20s.

The aim of the thesis was to produce an estimate of the Cost-effectiveness (CEA) and Cost-utility (CUA) of two alternative timings for the initiation of LTMV in three identical cohorts of 100 hypothetical individuals with DMD. The alternative timings were: (i) from the onset of the sleep disordered breathing (SDB) in rapid eye movement (REM) sleep, without hypercapnia (Intervention 1) and; (ii) from the onset of SDB in REM sleep, with hypercapnia (Intervention 2). The comparator was the current practice initiation of LTMV from the onset of diurnal hypercapnia.

In the base-case analysis of the CEA (all costs and outcomes discounted by 3%) Intervention 1 and 2 were dominated by the comparator. The probability (for a maximum acceptable ratio of 100,000 a life-year) that the comparator would be cost-effective was approximately 0.50 to 0.74; for Intervention 2 it was approximately 0.28 to 0.33 and; for Intervention 1 it was 0.04 to 0.18. In the base-case of the CUA (all costs and outcomes discounted by 3%) there were no strategies that were clearly dominated by any other. The probability (for a maximum acceptable ratio of $100,000 a QALY), that the comparator would be cost-effective was approximately 0.35 to 0.52; for Intervention 2 it was approximately 0.29 to 0.33 and; for Intervention 1 it was approximately 0.17 to 0.42.

The limited evidence for the effectiveness of the intervention, the inability of the small scale study to recruit and retain a statistically significant sample and the use of a researcher developed model of LTMV service delivery, mean that the results of the economic evaluation cannot be seen to be anything other than exploratory in nature.
Declaration

This is to certify that

(i) the thesis comprises only my original work towards the PhD except where indicated in the preface,
(ii) due acknowledgement has been made in the text to all other material used,
(iii) the thesis is less than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices

Stephen James Colgan
Acknowledgements

My supervisory panel, Professor Mark Jenkins, Professor Andrew Kornberg, Professor Mark Cook and Steven Crowley.

Dr Louisa Flanders, Associate Professor Jane Mosely and Professor Edward Byrne for their help and support.

Dr. Rosetta Marotta, Dr. Sharon Wong, Dr. Anita Quigley and all the staff at the National Muscular Dystrophy Research Institute, St Vincent's Hospital, Melbourne.

Boris M Struk, Muscular Dystrophy Association of Victoria.

Dr. K.Y. Tang for his help and support over the years.

Margaret Barlow for her careful and detailed proof reading.

Paula Barlow, for all the love and support she has given me throughout all my years of study.
# Table of contents

ACKNOWLEDGEMENTS......................................................................................................................... 7

TABLE OF CONTENTS....................................................................................................................................... 9

LIST OF TABLES............................................................................................................................................... 13

TABLE OF FIGURES......................................................................................................................................... 15

CHAPTER 1: INTRODUCTION.......................................................................................................................... 17

  BACKGROUND TO THE STUDY ......................................................................................................................... 17
  AIM OF THE THESIS ..................................................................................................................................... 21
  STRUCTURE OF THE THESIS .......................................................................................................................... 23

CHAPTER 2: DUCHENNE MUSCULAR DYSTROPHY ......................................................................................... 27

  INTRODUCTION ............................................................................................................................................ 27
    HISTORY AND CHARACTERIZATION ............................................................................................................ 28
    DIAGNOSIS ................................................................................................................................................ 29
    INHERITANCE ........................................................................................................................................... 30
    EPIDEMIOLOGY ......................................................................................................................................... 35
    NATURAL HISTORY .................................................................................................................................... 36
    CLINICAL MANAGEMENT .......................................................................................................................... 42
    ALTERNATIVE TREATMENT PATHWAYS .................................................................................................... 48

CONCLUSION.................................................................................................................................................. 51

CHAPTER 3: THE ECONOMIC EVALUATION OF HEALTHCARE INTERVENTIONS ..................................... 53

  INTRODUCTION ............................................................................................................................................ 53
    WELFARE ECONOMICS ............................................................................................................................... 54
    HEALTHCARE AS A COMMODITY ............................................................................................................... 57
    ECONOMIC EVALUATION ............................................................................................................................ 59
    COST-BENEFIT ANALYSIS (CBA) ............................................................................................................... 60
    COST-EFFECTIVENESS ANALYSIS (CEA) .................................................................................................. 63
    COST-UTILITY ANALYSIS (CUA) ............................................................................................................... 65
    THE DESIGN AND ANALYSIS OF ECONOMIC EVALUATIONS .................................................................. 69
    STUDY QUESTION ....................................................................................................................................... 69
    SELECTION OF THE INTERVENTION AND COMPARATOR ......................................................................... 70
    EVENT PATHWAY ....................................................................................................................................... 70
List of tables

Table 1: Clinical and diagnostic criteria for DMD

Table 2: Criteria to be met for a definite diagnosis

Table 3: Vital capacity and total lung capacity in relation to age

Table 4: Maximal respiratory pressures in relation to age

Table 5: Abnormal respiratory events during sleep

Table 6: Clinical management from the loss of ambulation, prior to & after the initiation of LTMV

Table 7: Multi-attribute utility scales

Table 8: Search strategy used in Medline, Cinahl & Embase

Table 9: Search strategy used in OpenSigle

Table 10: Excluded studies - Long term mechanical ventilation (n=140)

Table 11: Excluded studies - Quality of life in LTMV (n=27)

Table 12: Summary of the included studies (n=4)

Table 13: Methodology quality of the included studies (n=4)

Table 14: Other data collected for the decision-analytic model

Table 15: Prevalence of LTMV in Europe

Table 16: Reimbursement for LTMV in France

Table 17: VRSS LTMV service delivery model

Table 18: Model of LTMV service delivery that was used in the economic evaluation

Table 19: Capital equipment & the supply and replacement of consumables

Table 20: Number of questionnaires returned by study participants for each mail out (n=12)

Table 21: Individual results for the MRF28 (1 of 2)

Table 22: Individual results for the MRF28 (2 of 2)

Table 23: Individual results for the AQoL1

Table 24: Resource usage for the interventions, and the comparator

Table 25: Data sources for the valuation of resource usage

Table 26: Resource usage, unit costs and data sources

Table 27: Proportions of death by cause, proportion ventilated by cause and HRQoL

Table 28: Life years spent without and with LTMV

Table 29: Mean annual cost for 100 individuals (undiscounted)

Table 30: Summary of results - Cost-effectiveness analysis (Life years)

Table 31: Summary of results - Cost-utility analysis (QALYs)

Table 32: Multivariate analysis - CP_NoSCM_YV & IntV2_NoSCM_YV (Life years)

Table 33: Multivariate analysis - AQoL_V & CP_NOSCM_VP (QALYs)
Table 34: Multivariate analysis - AQoL_V & IntV1_NoSCM_VP (QALYs) ........................................228
Table 35: Multivariate analysis - AQoL_V & IntV2_NoSCM_VP (QALYs) ........................................230
Table 36: Multivariate analysis - CP_NoSCM_VP & IntV1_NoSCM_VP (QALYs) ...............................232
Table 37: Multivariate analysis - CP_NoSCM_VP versus IntV2_NoSCM_VP (QALYs) .....................234
Table 38: Multivariate analysis - IntV1_NoSCM_VP versus IntV2_NoSCM_VP (QALYs) ...............236
Table 39: Probabilistic analysis - Summary of results (Life years) .............................................238
Table 40: Probabilistic analysis – Summary of results (QALYs) ..................................................242
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Cost-effectiveness plane</td>
<td>90</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Example of a decision analytic model</td>
<td>103</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Example of a Markov model</td>
<td>108</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Quorum flow chart</td>
<td>130</td>
</tr>
<tr>
<td>Figure 5</td>
<td>The decision analytic model</td>
<td>169</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Condition specific quality adjusted life expectancy (Life years)</td>
<td>212</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Utility weighted quality adjusted life years (QALYs)</td>
<td>215</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Univariate analysis - Tornado diagram (Life years)</td>
<td>217</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Univariate analysis - Tornado diagram (QALYs)</td>
<td>219</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Multivariate analysis - CP_NoSCM_YV &amp; IntV2_NoSCM_YV (Life years)</td>
<td>222</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Multivariate analysis - AQoL_V &amp; CP_NoSCM_VP (QALYs)</td>
<td>225</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Multivariate analysis - AQoL_V &amp; IntV1_NoSCM_VP (QALYs)</td>
<td>227</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Multivariate analysis - AQoL_V &amp; IntV2_NoSCM_VP (QALYs)</td>
<td>229</td>
</tr>
<tr>
<td>Figure 14</td>
<td>Multivariate analysis - CP_NoSCM_VP &amp; IntV1_NoSCM_VP (QALYs)</td>
<td>231</td>
</tr>
<tr>
<td>Figure 15</td>
<td>Multivariate analysis - CP_NoSCM_VP &amp; IntV2_NoSCM_VP (QALYs)</td>
<td>233</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Multivariate analysis - IntV1_NoSCM_VP &amp; IntV2_NoSCM_VP (QALYs)</td>
<td>235</td>
</tr>
<tr>
<td>Figure 17</td>
<td>Probabilistic analysis - Intervention 2, scatter plot of results (Life years)</td>
<td>239</td>
</tr>
<tr>
<td>Figure 18</td>
<td>Probabilistic analysis - Intervention 1, scatter plot of results (Life years)</td>
<td>240</td>
</tr>
<tr>
<td>Figure 19</td>
<td>Cost-effectiveness acceptability curve (Life years)</td>
<td>241</td>
</tr>
<tr>
<td>Figure 20</td>
<td>Probabilistic analysis - Intervention 2, scatter plot of results (QALYs)</td>
<td>243</td>
</tr>
<tr>
<td>Figure 21</td>
<td>Probabilistic analysis - Intervention 1, scatter plot of results (QALYs)</td>
<td>244</td>
</tr>
<tr>
<td>Figure 22</td>
<td>Cost-effectiveness acceptability curve (QALYs)</td>
<td>245</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

Background to the study

Duchenne muscular dystrophy (DMD) is an X-linked recessive and invariably fatal neuromuscular disease that affects approximately one in every 3300 live male births. The underlying cause of the disease is a mutation of the dystrophin gene and the resultant inability to produce a protein known as dystrophin. One of the functions of this protein is to stabilise and protect the plasma membranes of muscle cells from the mechanical stresses induced by muscle fibre contraction. In the absence of this protein muscle contraction disrupts the cell membranes, triggering a process that ultimately ends in the death of the muscle cells and a reduction in muscle strength and function. The first symptoms of the disease are seen in early childhood, when symmetrical proximal hip and lower limb muscle weakness makes it difficult for affected boys to walk, run or climb stairs. As the child ages there is increasing involvement of the muscles of the upper limbs and neck. Cardiac muscle involvement, while sub-clinical in most cases until very late in the disease, is severe enough in 10-20% of individuals to result in death at a mean age of 16 years. The first signs of the decline of respiratory muscle strength and function are seen during sleep, which is disrupted by periods of abnormal breathing (Sleep-disordered breathing). Initially this only occurs during REM sleep but as respiratory function continues to decline it occurs throughout all of the stages of the sleep cycle. When vital capacity, a measure of respiratory function falls to less than 20% of predicted (for height, age and sex) adequate ventilation cannot be maintained and unless some form of long-term mechanical ventilation (LTMV) is used to support respiratory function it will invariably result in respiratory failure and death (Hughes et al, 1996; Bakker & Van Ommen, 1998; Sivak, Shefner & Sexton, 1999; Hoffman, 2001; Biggar et al, 2002; Birnkrant, 2002; Eagle et al, 2002; Piper, 2002; Muntoni, 2003; Finsterer & Stollberger, 2003).

There is currently no cure for DMD and consequently the aim of clinical management is the early detection and treatment of complications. In the early ambulant stages of the disease the mainstay of treatment is physiotherapy and the use of individualised orthoses to prevent or reduce the severity of contractures. While corticosteroids are often used alongside physiotherapy to maintain muscle strength and to delay the loss of ambulation and the development of scoliosis, there is no consensus on their use as a standard treatment. After the loss of ambulation the focus of management shifts to: (i) the surgical correction of scoliosis and knee and hip contractures; (ii) the provision of psychosocial support; (iii) the prevention of malnutrition and obesity; (iv) the monitoring of cardiac function and the treatment of cardiomyopathy and; (v) the monitoring of respiratory muscle strength and function and the support of respiratory function through the use of LTMV. The main cause of death in DMD is respiratory failure, and consequently the
early detection and treatment of respiratory complications is a major component of the management of the
disease. While the use of long-term mechanical ventilation to support respiratory function is
recommended in current practice guidelines, there is a lack of clear and consistent criteria for when the
intervention should be initiated. In individuals with DMD, LTMV can be initiated from either: (i) the onset of
sleep-disordered breathing in REM sleep without hypercapnia; (ii) the onset of sleep-disordered breathing
in REM sleep with hypercapnia or; (iii) shortly prior to, or after the onset of diurnal hypercapnia (when vital
capacity falls below 1000mls) (Emery & Muntoni, 2003; Manzur & Muntoni, 2002; Manzur et al, 2004;
Ramelli & Hammer, 2005; Toussaint, Chatwin & Soudon, 2007; Wagner, Lechtzin & Judge, 2007; Bushby
et al, 2010a; Bushby et al, 2010b).

While the use of LTMV to support respiratory function is recommended practice for individuals with DMD,
there are still many doctors who do not routinely offer it to their patients. A survey conducted by Gibson
(2001) of neuromuscular specialists in Canada, found that only 22.7% routinely raised the possibility of the
use of LTMV and that many of those who did framed their discussion in such a way as to effectively
discourage its use. The reasons provided by the specialists for doing this were: (i) that the quality of life of
a ventilator user was poor; (ii) that the cost of treatment was prohibitively high; (iii) that it placed an
excessive burden on families and; (iv) it was a subject that many families were uninterested in, or unwilling
found that: (i) 13% of respondents believed the quality of ventilator users was poor and; (ii) 34%
considered the use of LTMV for individuals with DMD to be an inappropriate use of public funds as it was
not a cost-effective treatment (Leger & Hill, 2001; Leger & Laier-Groeneveld, 2002; American Thoracic
Society, 2004; Ramelli & Hammer, 2005; Toussaint, Chatwin & Soudon, 2007; Muntoni, 2010).

There is actually very little known about the quality of life of individuals with DMD who have their
respiratory function supported by LTMV. The majority of the studies that have been conducted of the
quality of life of ventilator users are not DMD specific, and instead report the aggregate quality of life for: (i)
groups of individuals with a range of neuromuscular diseases, of which DMD is only one type; (ii) a mix of
different types of LTMV (i.e. negative pressure, invasive positive-pressure, non-invasive positive pressure)
and; (iii) the use of LTMV in a range of settings (i.e. at home, as an inpatient in an acute hospital, or at a
long-term care facility). There is also very little known about either the cost or Cost-effectiveness of the
use of LTMV for neuromuscular diseases and chest wall disorders. At the time this study was conducted
there were no economic evaluations of the use of LTMV for individuals with DMD, and consequently it is
not known whether its use in this disease group has significant resource use implications (Fauroux &
Dreyer, Steffensen & Pedersen, 2010a).
In the last twenty-five years there has been a rapid increase in the use of domiciliary LTMV to support respiratory function in individuals with neuromuscular diseases and chest wall disorders. This is due to an increase in the number of diseases and conditions in which LTMV is used and the commercial development of nasal masks and relatively inexpensive non-invasive, pressure cycled, positive pressure ventilators. The largest, longest running and best organised formal system for the provision of domiciliary LTMV is ANTADIR (Nationale pour le Traitement A Domicile de l'Insuffisance Respiratoire Chronique) in France, which in 1996 provided LTMV and home oxygen to over 35,000 adults and children at an annual cost of 92 million Euros. The only formal system in Australia for the provision of domiciliary LTMV is the state based Victorian Respiratory Support Service (VRSS). The VRSS was established in 1995 at the Austin and Repatriation Medical Centre (A&RMC), when a program run by the Fairfield Hospital for ventilator dependent polio survivors and a program run by the A&RMC for neuromuscular and chest wall diseases were combined into a single centralised service. In less than 10 years there was a 155% increase in the number of patients treated by the VRSS. In 2004 the VRSS provided LTMV to 434 adults, of which 155 (35.71%) were individuals with neuromuscular diseases and chest wall disorders, at a cost to government of $940,000 AU. The proportion of this cost that is attributable to the provision of domiciliary LTMV for individuals with DMD is not known, as there are no disease specific costs reported by the VRSS (Sakakihara et al, 1996; Health Solutions International, 1997; Campbell & Pierce, 1998; Donner et al, 2001; Leger & Laier-Groeneveld, 2002; d’Orbcastel, 2004; Stuart & Weinrich, 2004; Department of Human Services, 2005; Lloyd-Owen et al, 2005; Annane et al, 2009).

In recent years governments have come under increasing pressure to contain or reduce the cost of their healthcare systems. To attempt to do this decision makers have come to demand evidence of Cost-effectiveness as a key criterion on which to base new and recurrent funding decisions. Cost-effectiveness analysis, a form of economic evaluation, is an explicit non-market mechanism that was developed to generate information about the costs and benefits of alternative healthcare interventions that could be used by decision-makers to maximise the benefits gained from the use of their limited resources. In Australia and in the United Kingdom, The Netherlands, Canada and Scandinavia, economic evaluation is an essential component of the decision making process for government funding decisions for medical interventions, devices and pharmaceuticals (Richardson et al, 1998; Carter & Harris, 1999; George, Harris & Mitchell 2001; Leger & Laier-Groeneveld, 2002; Drummond, 2004; Drummond et al, 2005; Brodtkorb, Henriksson, Johannesen-Munk, 2008; Annane et al, 2009).
Aim of the thesis

The aim of the thesis is to produce an estimate of the Cost-effectiveness and Cost-utility of the use of domiciliary long-term non-invasive positive pressure ventilation to support respiratory function in individuals with Duchenne muscular dystrophy.

In the thesis the costs and outcomes of two alternative treatment pathways for the timing of the initiation of LTMV in DMD, specifically: (i) from the onset of SDB in REM, without hypercapnia at 15.5 years and; (ii) from the onset of SDB in REM, with hypercapnia at 17.5 years, were compared to its current practice initiation at 18.5 years of age. The scope of the thesis was limited to the use of non-invasive positive pressure LTMV. There was no consideration of the cost or effectiveness of other forms of LTMV, such as non-invasive negative pressure ventilation or invasive positive pressure ventilation, as these forms of mechanical ventilation are not generally used to support respiratory function in individuals with DMD.

Domiciliary non-invasive positive pressure LTMV is delivered via a small, lightweight clear plastic nasal mask with a soft silicone rubber cuff, which forms an air seal over the skin surrounding the nose. The mask is held in place by a set of wide elastic straps that fit around the back of the head, with a flexible, lightweight but rigid plastic tube connecting the mask to a positive pressure ventilator, a small lightweight microprocessor controlled device. When the ventilator detects a change in the pressure in the closed system formed by the mask and tubing, the machine augments the inhalation pressure generated by the individual, and fully inflates the lungs and reduces the muscular effort needed to breathe. At the end of inhalation, exhalation occurs simply as a result of the elastic recoil of the lungs with the ventilator producing only enough pressure to help prevent the collapse of the upper airways, and to vent carbon dioxide from the mask and tubing (Mehta & Hill, 2001).

There is currently no cure for DMD, and the only intervention that has been shown to have any significant impact on life expectancy is the use of LTMV to support respiratory function. Consequently the use of LTMV in DMD is an important area of study. In the last twenty years research into rare diseases, such as DMD, has been increasingly recognised as an important medical, social and public-health issue, as these diseases affect more than 50 million people worldwide. It is only through research that these individuals can hope to gain equitable access to effective healthcare interventions. The vulnerable nature of individuals with rare diseases, most of whom are children, raises ethical issues that can make it difficult to conduct research studies in these populations. The limited number of individuals with these diseases makes it difficult to recruit and retain the number of study participants needed to produce a statistically significant result. Unfortunately, as a result of these factors, the best available evidence for the
effectiveness of an intervention for a rare disease may consist of nothing more than a limited number of uncontrolled studies with very small sample sizes (Eagle et al, 2002; Schieppati et al, 2008; Panju & Bell, 2010; Gliklich & Leavy, 2011).

These and other factors make it a challenge to conduct an economic evaluation of the use of domiciliary non-invasive LTMV in DMD, as to do so requires a detailed understanding of: (i) the disease, its natural history and management; (ii) the decline of respiratory function and how this is monitored in clinical practice; (iii) the development and staging of sleep disordered breathing and how it is detected; (iv) the different points in the decline of respiratory function at which LTMV can be initiated; (v) the evidence for effectiveness of the use of the intervention; (vi) the services that are currently in place for the provision of domiciliary LTMV in Australia and overseas and (vii); the machines and consumables that are needed to provide the intervention.

The thesis aims to add to existing knowledge by undertaking and reporting the findings of:

1. A formal systematic literature review of the best available evidence, suitable for use in an economic evaluation, of the effectiveness of domiciliary LTMV for individuals with DMD;
2. A small scale study of the condition specific and generic utility weighted quality of life of non-ambulant individuals with DMD, prior to and after the initiation of LTMV, to collect data suitable for use in the economic evaluation and;  
3. An economic evaluation of the use of domiciliary LTMV to support respiratory function in DMD that combines, in a decision analytic model, the evidence for the effectiveness of the intervention obtained from the systematic literature review and the estimates of quality of life obtained from the small scale study.
Structure of the thesis

The first chapter of the main body of the thesis, Chapter 2, contains a detailed overview of the disease, its history and characterisation, how it is diagnosed, its mode of inheritance, epidemiology, natural history and management in current practice. The focus of the chapter is on the monitoring and early detection of the decline of respiratory function, the use of LTMV and the details of the alternative treatment pathways that were examined in the economic evaluation.

The economic evaluation of healthcare interventions is underpinned by Welfare economic theory; an analytical framework for assessing the desirability of the alternative allocation of resources for the production and distribution of goods and services, with the aim of maximising social welfare through the efficient allocation of resources. There are a number of key assumptions within this theory: (i) individuals are the best judge of their own welfare and act in a rational and logically consistent manner to maximise their utility; (ii) individuals maximise their utility through, among other things, their consumption of goods and services in a free and perfectly competitive market and; (iii) the overall welfare of society can be determined from the sum of individual utility. Furthermore, it is assumed that there is complete certainty in a perfectly competitive market, with no random events and information is freely available to all. The mix of goods and services, and the quantities produced within this market are exactly equal to consumer demand, with the prices paid by consumers and the payments received by producers solely the result of market forces (the independence of supply and demand). There is supply side competition and due to the presence of competitive equilibrium, these goods and services are produced in a technically efficient manner, with the prices charged by producers accurately reflecting the opportunity cost, the mix of labour and capital used in their production and the return that could be have been gained from the alternative use. If there are no increasing returns from production and every good or service an individual may wish to consume is priced and available in this market, then this will result in an allocation of resources that is Pareto optimal for any given distribution of purchasing power, as there is no other allocation of resources that would improve the utility of one individual, without decreasing the utility of another. A market or commodity that deviates in any significant way from the underlying assumptions of this model of perfect competition will be unable to achieve optimality.

Unfortunately the characteristics of health care as a commodity are far removed from those required for a perfectly competitive market. For example, there is significant uncertainty surrounding the decisions that a consumer needs to make about their purchase of healthcare in relation to their risk of disease, their need for treatment, and the effectiveness of treatment. A consumer who purchases the services of a medical or other health professional does this, to some extent, to attempt to reduce this uncertainty through the
purchase of knowledge. In doing so the consumer, may find that based on information that is only available to the health professional that they may be prescribed treatments for problems that they were unaware they had, which while being appropriate, may be expensive, invasive or dangerous. It is generally accepted that these and other factors, such as the presence of significant externalities, have resulted in pervasive market failure and the inefficient allocation of health care resources and a resultant need for the explicit evaluation of the costs and outcomes of alternative healthcare interventions. The welfare economic framework, and the characteristics of healthcare as a commodity are described in detail in Chapter 3, and there is an introduction to the methods that are used in the design, analysis and reporting of the economic evaluation of healthcare interventions.

Decision analytic modelling is a systematic framework that is used to structure imperfect information about events and outcomes in situations where there is significant uncertainty and multiple competing objectives, and a clear need to determine the option that returns the highest expected value. It is an approach that is widely used in the economic evaluation of healthcare interventions to synthesize the best available evidence about the incidence and natural history of a disease, the effectiveness of at least two alternative treatment options, the impact that these have on outcomes such as life expectancy and quality of life and disease related events and activities that consume healthcare resources. In Chapter 4 there is a description of the process of developing a decision analytic model for use in an economic evaluation, and how the choice of the type of model and its structure are driven to a major extent by the characteristics of the disease and intervention under evaluation, characteristics such as: (i) the nature of the disease, i.e. chronic as opposed to acute; (ii) the health-related events that may occur over time; (iii) the nature of the risk of these health-related events, and whether this is constant, or varying over time and; (iv) whether the risk of the occurrence of future health related events are determined by the previous health related events experienced by an individual.

Mechanical ventilation is the application of pressure, to the body or to the airways, to support or replace inspiratory and expiratory muscle function. The earliest mechanical ventilators were large, heavy non-invasive negative pressure devices that intermittently reduced the air pressure over the chest wall to less than atmospheric pressure to create a pressure differential to inflate the lungs. While the use of negative pressure ventilation clearly saved the lives of many patients with acute respiratory muscle paralysis during the poliomyelitis epidemics of the late 1930-1950s, only 1 in 5 survived. In 1952, during a polio epidemic in Copenhagen, a shortage of negative pressure ventilators saw some individuals treated with invasive positive pressure devices which were normally only used during surgery. There was an unexpected reduction in mortality of 50% associated the use of these devices, which resulted in a gradual move towards the use of invasive positive pressure ventilation for the treatment of short term and long term
respiratory muscle failure. The high cost of invasive positive pressure ventilation and the complications associated with the need for a tracheotomy, ultimately limited its usefulness as a form of LTMV. Negative pressure ventilation remained the most widely used form of LTMV until the early 1980s, when it was replaced by non-invasive positive pressure ventilation. In Chapter 5 the evidence for the effectiveness of the use of non-invasive positive pressure domiciliary LTMV to support respiratory function for individuals with DMD was examined in a formal systematic literature review.

The first domiciliary LTMV services were developed in France to care for the ventilator dependent survivors of the polio epidemics of the 1940s and 1950s. Since that time these services have evolved to provide LTMV to individuals with a wide range of neuromuscular diseases and chest wall disorders. In Europe alone, there are now 329 centres across 16 countries for the provision of domiciliary LTMV. In Australia the delivery of LTMV is undertaken mainly on an ad-hoc basis by the respiratory departments of major tertiary referral hospitals. The cost of providing domiciliary LTMV in Australia for individuals with DMD is not known, and it was therefore necessary in Chapter 6 to construct a model of LTMV service delivery that could be as a basis of a cost estimate for use in the economic evaluation.

The methods that were in the economic evaluation to combine the events in the clinical treatment pathways, with the evidence of the effectiveness of the intervention and the equipment and staffing requirements of a service for the provision of domiciliary LTMV are described in detail in Chapter 7. The systematic literature review was unable to locate any studies of the health related quality of life (HRQoL) of individuals with DMD that were suitable for use in the economic evaluation. It was therefore necessary to undertake a small study to attempt to collect this information. Unfortunately the limited evidence for the effectiveness of the intervention, the inability of the HRQoL study to recruit and retain a statistically significant sample and the use of a researcher developed LTMV service delivery model mean that the results of the evaluation cannot be seen to be anything other than exploratory in nature.

The results of the base-case Cost-effectiveness and Cost-utility analysis and the findings of univariate, multivariate and probabilistic sensitivity analyses are reported in Chapter 8. In Chapter 9, there is a summary of the results of the evaluation and a guide to their interpretation. There is also a discussion of the limitations of the evaluation and the generalisability and transferability of the results. The results of the evaluation are compared to the findings of studies of the use of LTMV in Amyotrophic lateral sclerosis (Motor neuron disease). The distributive implications of economic evaluations and the trade-off between efficiency and equity are then examined, along with the potential for double jeopardy in the strict application of the QALY framework. The implications of the public preference for interventions that can avert or delay the death of seriously ill individuals are then examined and the development, in Australia, of
public policy frameworks that explicitly take this into account. The need for future research is discussed and how the difficulties in obtaining a statistically significant sample in studies conducted in rare disease populations could be addressed through the use of innovative research designs and bio statistical techniques and the use of disease registries. The chapter ends with a discussion of how the many limitations of the evaluation could be addressed through further research, such as: (i) a replication and extension of the systematic literature review, conducted by two researchers, that included full professional translations of foreign language studies and; (ii) an Australia-wide, multi-institution longitudinal prospective study of the HRQoL of individuals with a DNA confirmed diagnosis of DMD.
Chapter 2: Duchenne muscular dystrophy

Introduction

This chapter provides the reader with an introduction to Duchenne muscular dystrophy, the natural history of the disease, how it is currently managed and the alternative treatment pathways for the timing of the initiation of long term mechanical ventilation (LTMV). The main focus of the chapter is on the decline of respiratory function in the disease, the development of sleep-disordered breathing, and the use of LTMV.

The chapter begins with a brief history of the disease, from the early work of Richard Partridge and Edward Meryon, to that of Guilluame-Benjamin-Amand Duchenne. This is followed by an overview of the presentation and diagnosis of the disease in early childhood and the discovery in 1982 of the genetic mutation of the dystrophin gene that is the underlying cause of the disease. The impact that the lack of dystrophin, the protein that is normally produced by this gene, has on muscle function and cognition is then examined. The first section of the chapter ends with an overview of the epidemiology of the disease and what is known about its natural history, with particular attention paid to the decline of respiratory muscle strength and function and the development of sleep-disordered breathing.

The second half of the chapter begins with a detailed description of the current management of the disease, starting with the use of physical therapy and corticosteroids to prevent or delay the development of contractures and scoliosis, and the need for corrective surgery. This followed by a brief overview of: the monitoring of cardiac function and the management of cardiac failure; the management of malnutrition and obesity and; the provision of psychosocial support. The monitoring and management of the decline of respiratory muscle strength and function is then examined in detail. The chapter ends with details of the four distinct stages in the decline of respiratory function in the disease and the three distinct alternative pathways for the timing of the initiation of LTMV in DMD that were examined in the economic evaluation.
History and characterization

Duchenne muscular dystrophy (DMD) was the first well-characterized form of muscular dystrophy, due to it being a relatively common disease that was associated with an easily recognized paradoxical enlargement of severely weakened muscles, a feature known as “pseudo-hypertrophy”. The first clinical and pathological reports of the disease were made in 1847, by Richard Partridge an English physician. While Partridge’s description of the clinical features of the disease were brief, and lacked sufficient detail to firmly establish a diagnosis, the patient he described and his affected siblings were later reported in more detail by Edward Meryon (Emery & Muntoni, 2003; Tyler, 2003).

Edward Meryon was an English surgeon who studied medicine in Paris and London. In 1851 Meryon submitted a detailed report to the Royal Medical and Chirurgical Society of London of the first systematic clinical, genetic and pathological description of the disease. In this report, which was published in 1852, Meryon described the clinical presentation of the disease, from its onset in early childhood to death in late adolescence. Meryon described eight cases of the disease in three different families and its affinity for males. The availability of pathological material from the autopsy of two boys meant that Meryon, a founding member of the Microscopical Society of London, was able to provide details of the microscopic examination of muscle and other tissue. Meryon reported that the major abnormalities of the disease were seen in skeletal muscle which he described as being soft, atrophied, and almost bloodless and lacking the deep red color of normal muscle. Meryon also described the destruction of striped elemental fibres of the muscle and the muscle tunic (sarcolemma) and their replacement with granular matter and oil globules (Emery, 1993; Hoffman, 2001; Emery, 2002a; Emery & Muntoni, 2003; Tyler, 2003).

Guillaume Benjamin Amand Duchenne was born in 1806 and studied medicine in Paris. His study of the disease began in 1858 when his attention was drawn to a case in his private clinic which he briefly described in a book first published in 1861. In this book Duchenne described a 9-year-old boy with the classic symptoms of the disease, delayed motor milestones, calf muscle hypertrophy, an inability to stand from a sitting position without the use of the hands, and general difficulty ambulating. Between 1861 and 1868 Duchenne systematically and painstakingly described the clinical, genetic and pathological features of the disease that now bears his name. Duchenne unlike Meryon, who only had access to tissue samples from autopsy, was able to study the histology of affected muscles in the same patient at different stages of the disease with relative safety, through the examination of muscle biopsy samples obtained from live individuals. These samples were obtained using a needle harpoon (enporte-piece histologique) that was developed by Duchenne. Prior to the development of his “needle harpoon” the only method for obtaining tissue samples was from an open surgical biopsy, a difficult procedure that was not without risk,
as it required the use of a general anaesthetic, and recovery was often complicated by postoperative pain and wound infections that could take weeks to clear. While Duchenne initially referred to the disease as hypertrophic paraplegia of infancy, in later years he referred to it as pseudohypertrophic muscular paralysis (paralysie musculaire pseudohypertrohique) and paralytic sclerosis of muscle (paralysie myosclerosique) to better reflect the natural history of the disease and his evolving understanding of it. Duchenne ultimately characterised the disease as a form of progressive muscle weakness, first of the lower limbs, then later of the upper limbs that was accompanied by a gradual increase in the size of affected muscles as a result of an increase in interstitial, fibrous and adipose tissue (Emery, 1993; Hoffman, 2001; Emery & Muntoni, 2003; Tyler, 2003).

**Diagnosis**

A preliminary diagnosis of DMD is based on clinical presentation, with the onset of the disease usually apparent before three years of age. The most common presentation of the disease is a delay in the age at which the child begins to walk, although it can present as a non-motor developmental disorder in which speech is delayed. The mean age at which children normally learn to walk is approximately 13 months with only 3% not walking by 2 years. In a case series of 114 children with DMD, 56% took at least 18 months before they learnt to walk, with a further 25% taking at least 24 months. The delay is the result of hip and proximal limb weakness that makes it difficult for children with DMD to walk with other than an unsteady, broad based, waddling gait. There is a tendency to walk on tiptoes, a symptom that worsens when they are tired, and there is difficulty climbing stairs and a tendency to fall. In the early stages of the disease the lower limbs are affected more than the upper limbs, with proximal muscles affected more severely than the distal muscles. Muscle involvement is bilateral and symmetrical with a differential pattern of involvement that results in one of the characteristic physical features of the disease, that is known as “Gower’s sign”, where to stand from lying on the floor children with DMD need to use their arms to push themselves into a standing position, effectively climbing up their thighs to extend their hips and push out their trunk (Bakker & Van Ommen, 1998; Mohamed, Appleton & Nicolaides, 2000; Manzur & Muntoni, 2002; Emery & Muntoni, 2003).
Inheritance

It was noted very early in the history of the disease that DMD was transmitted by healthy females to their male offspring, a mode of inheritance known as X-linked recessive\(^1\). It is now known that the disease is inherited in only 30% of cases, with the majority of cases being the result of spontaneous genetic mutations in the males themselves, or in the ova of their mothers. In 1982 the gene responsible for DMD, the dystrophin gene, was located on the short arm of the X-chromosome (XP-21) (Emery, 2002b). This gene is one of the largest in the human body and contains 2.7 million base pairs that encode a total of 79 exons\(^2\). The size of the gene makes it extremely vulnerable to spontaneous rearrangement and recombination events and the consequent production of the genetic mutations that are the underlying cause of the disease. The expression of the gene in the cells of the body is controlled by three independent promoters, B, M and P. The B promoter is found in cortical neurons and the hippocampus of the brain and produces only low levels of the dystrophin protein. The M promoter in skeletal and cardiac muscle cells produces high levels of the protein, as does the P promoter in the cerebellum (Peterlin et al, 1997; Wilton et al, 1997; Bakker & Van Ommen, 1998; Ozcelik, 2002; Emery, 2002b; Emery & Muntoni, 2003; Kapsa, Kornberg & Byrne, 2003; Muntoni, Torelli & Ferlini, 2003; van Deutekom & van Ommen, 2003).

The product of the dystrophin gene, in muscle cells (myofibers), is a cytoskeletal protein that localizes to the sarcolemma, the sheath that surrounds the myofibers. The protein provides a crucial structural element that flexibly anchors the actively contracting component of the muscle, the intracellular Actin cytoskeleton, across the sarcolemma to the extracellular matrix. This acts to stabilize and protect the sarcolemma from the mechanical stresses induced by muscle fibre contraction. In the absence of this protein muscle contraction disrupts the sarcolemma producing an influx of extracellular calcium into the myofibers that triggers a process that leads to cell autolysis and myofiber necrosis. Myofibers are highly specialised cells and one consequence of this specialization is the loss of the ability to replicate. Myofiber growth, repair and replacement is instead reliant upon a finite population of muscle precursor cells, a type of adult stem cell known as satellite cells, that are found within muscles. In DMD ongoing contraction induced myofiber damage significantly increases the demands placed upon these precursor cells for muscle repair and replacement and they are rapidly depleted. When the supplies of satellite cells are exhausted the myofibers die and are replaced with fat and fibrous tissue, with a consequent reduction in

---

\(^1\) X-linked recessive diseases are inherited from genes on the two chromosomes that determine sex. The Y chromosome inherited from the male is 1/3 the size of the X chromosome inherited from the female, and carries approximately 15 genes, with the X chromosome carrying 2500. When a male inherits a recessive gene from his mother, the trait encoded for by that gene is always expressed, and not damped or masked as with other recessive traits, due to there being no corresponding gene on the Y chromosome. For a female to express an X linked recessive trait it must be present on both of the X chromosomes inherited from her parents (Marieb, 1998).

A robust diagnosis of DMD is made on the basis of an accurate clinical assessment, elevated serum creatinine kinase levels, and DNA analysis of blood leukocytes. If DNA analysis is unable to detect the commonly occurring deletions that are the underlying cause of the majority of the cases of the disease, then, where possible gene sequencing should be used to search for the other deletions, duplications, insertions and point mutations that are known to cause DMD (Bakker & Van Ommen, 1998; Muntoni, 2001; Jay & Vajsar, 2001; Byrant-Greenwood, 2002; American Thoracic Society, 2004; Bushby et al, 2010a).

The diagnostic criteria for DMD from Bakker et al (1997) are listed in Tables 1 and 2.

---

2 Sequences of DNA in genes that encode information (Marieb, 1998).
Table 1: Clinical and diagnostic criteria for DMD

<table>
<thead>
<tr>
<th>#</th>
<th>Diagnostic criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Symptoms are present before the age of five years.</td>
</tr>
<tr>
<td>2</td>
<td>Progressive symmetrical muscle weakness; proximal limb muscles are affected more severely than distal muscles; initially only the lower limbs are affected; in many cases there is hypertrophy of the calf muscles.</td>
</tr>
<tr>
<td>3</td>
<td>No evidence of fasciculation or a loss of sensation.</td>
</tr>
<tr>
<td>4</td>
<td>Loss of unassisted ambulation before the age of 13 years.</td>
</tr>
<tr>
<td>5</td>
<td>Minimum of a 10-fold increase in serum creatinine kinase levels, relative to age and mobility.</td>
</tr>
<tr>
<td>6</td>
<td>If a muscle biopsy is performed, histopathological examination reveals abnormal variation in the diameter of muscle fibres, with atrophic and hypertrophic fibres, areas of necrotic and regenerative tissue and hyaline fibres and an increase of endomysial connective tissue and fat.</td>
</tr>
<tr>
<td>7</td>
<td>Less than 5% of dystrophin is expressed in stained tissue taken from muscle biopsy (immunochemistry); and there are occasional revertant fibres.</td>
</tr>
<tr>
<td>8</td>
<td>DNA analysis reveals a Duchenne type mutation within the dystrophin gene, or an identical haplotype, involving closely linked markers as in previous cases in family members</td>
</tr>
<tr>
<td>9</td>
<td>A positive family history of the disease; compatible with a X-linked recessive disease.</td>
</tr>
</tbody>
</table>

**Table 2: Criteria to be met for a definite diagnosis**

<table>
<thead>
<tr>
<th>#</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>For the first case in a family the following criteria must be met</td>
</tr>
<tr>
<td></td>
<td>(a) The patient is less than 5 years old and 2, 3, 5, 6, 7 &amp; 8 all are present.</td>
</tr>
<tr>
<td></td>
<td>(b) The patient is aged between 5 and 12 years and 1, 2, 3, 4, 5 (at least once), 6, 7 &amp; 8 are all present</td>
</tr>
<tr>
<td></td>
<td>(c) The patient is older than 12 years and 1, 2, 3, 4, 5 (at least once), 8, (or 6 &amp; 7) are all present</td>
</tr>
<tr>
<td>B</td>
<td>There is another case in the family (as per criteria 9) that meets all of the criteria of A.</td>
</tr>
<tr>
<td></td>
<td>(a) The patient is less than 5 years of age and 5 &amp; 9 are present</td>
</tr>
<tr>
<td></td>
<td>(b) The patient is aged between 5 and 12 years and 1, 2, 3 &amp; 5 (at least once) are all present</td>
</tr>
<tr>
<td></td>
<td>(c) The patient is older than 12 years of age and 1, 2, 3, 4, 5 (at least once) are all present</td>
</tr>
<tr>
<td>C</td>
<td>The diagnosis is possible when</td>
</tr>
<tr>
<td></td>
<td>(a) The patient is less than 5 years of age and 2, 3, 5, 6 are all present</td>
</tr>
<tr>
<td></td>
<td>(b) The patient is aged between 5 and 12 years and 1, 2, 3, 4, 5 (at least once), 6 are all present</td>
</tr>
</tbody>
</table>

Until recently few clinicians doubted their ability to make an accurate diagnosis of DMD, as it was considered one of the most stereotyped of all inherited human diseases. This has been shown not to be true, for in approximately 30% of cases of suspected disease the diagnosis cannot be confirmed by the DNA tests that are routinely used in clinical practice. Without confirmation of the diagnosis from DNA testing it can be difficult to differentiate DMD from other forms of childhood onset muscular dystrophy, such as Becker muscular dystrophy or Limb Girdle muscular dystrophy (Type 21) (Bushby, 2004; Griggs & Bushby, 2005; Schwartz et al, 2005).

Becker muscular dystrophy (BMD) was first described by Becker and Kiener in 1955. In general it is a less severe form of muscular dystrophy, but one that in its severest form can be clinically indistinguishable from DMD. Duchenne muscular dystrophy and Becker muscular dystrophy are both caused by mutations of the dystrophin gene. The mutations of the dystrophin gene that result in DMD disrupt the reading frame necessary for the production of the dystrophin protein, and there is no protein produced. The mutations that result in BMD leave the reading frame intact, and there is protein production but only in an abnormally small and partially functional form. Limb girdle muscular dystrophy (Type 21) is generally less severe than DMD and in the majority of cases the disease follows a mild Becker-like course. The disease has many clinical features in common with DMD, such as proximal muscle weakness, hypertrophy of calf muscles and elevated creatinine kinase levels and in severe cases it can follow a rapidly progressive Duchenne-like course, with an early loss of ambulation and respiratory failure in early adulthood (Wilton et al, 1997; Bakker & van Ommen, 1998; Yazaki et al, 1999; Blake et al, 2002; Emery & Muntoni, 2003; Schwartz et al, 2005).
Epidemiology

The majority of studies of the epidemiology of DMD were conducted prior to the discovery of the dystrophin gene. In these studies the incidence of the disease was reported to be 21 per 100,000 live male births, or approximately one in every 4,854. On the basis of a series of studies in which the diagnosis of the disease was confirmed by DNA testing, the incidence is now believed to be closer to 30 per 100,000 live male births, or approximately one in every 3,300. Although DMD is a lethal X-linked recessive disease that presents before reproductive age, its incidence remains stable in most populations due to a high rate of spontaneous mutations. The mean prevalence of the disease as reported by Emery & Muntoni (2003) is approximately 6 per 100,000 of the total male population. Numerous studies of the prevalence of DMD have been performed throughout the world, and the reported prevalence varies from between 2 and 9 per 100,000 of the total male population. The wide variation in the reported prevalence of the disease may be due to: (i) methodological differences in the studies, such as the use of different diagnostic criteria and methods of ascertainment; (ii) differences in the frequency of mutations across geographic areas or ethnic groups; (iii) the manipulation of the sex ratio at birth³, resulting in an abnormally high proportion of males in the population. In recent years studies of the frequency and pattern of genetic deletions in DMD have shown that there are statistically significant differences (chi-squared, p<0.001) in the distribution of deletions between population groups in Europe (Belgium, Denmark, Estonia, France, Hungary, Italy & the Netherlands), Arabia (Saudi Arabia and Egypt) and Southeast Asia (China, The Philippines & Japan) (Emery, 1991; Peterlin et al, 1997; Bakker & Van Ommen, 1998; Bushby et al, 2001; Emery, 2002a; Emery & Muntoni, 2003; Griggs & Bushby, 2005; Hesketh & Xing, 2006; Basak et al, 2009).

The manipulation of the sex ratio at birth in some parts of world (East Asia, South Asia, North Africa & The Middle east), due to a preference for male offspring, has resulted in a substantial distortion in the population sex ratio⁴. This preference for sons is manifested in practices such as prenatal sex determination by second trimester ultrasound and sex selective abortion, and the neglect or abandonment (and death) of female infants and children. In human populations there are normally between 105-107 male births for every 100 female births (a median of 105.9) and generally 50.4% of the total population is male, due to differences in gender specific age related mortality. The largest distortions of the population sex ratio are seen in Pakistan, India and Bangladesh, with the United Nations estimating that in 2001 there

³ The number of male live births for every 100 female live births. The accuracy of this figure varies between countries, due to differences in the registration of birth statistics, such as the incomplete recording of home births in countries with large rural populations, such as in India (Hesketh & Xing, 2006).

⁴ The number of males for every 100 females. This figure is generally seen to be more reliable than the sex ratio at birth, as it is generally taken from census data (Hesketh & Xing, 2006).
were 44 million less females in the population of India (93 females for every 100 males) than would be normally expected. While a law was implemented in India in 1996 that made pre-natal sex selection and abortion illegal (The Pre-natal Diagnostic Techniques - Regulation and Prevention of Misuse act) there is little evidence that it has had any real impact, as the population sex ratio in India in 2005 of 86/100 (95% CI: 81 to 92) was not significantly different from that reported prior to the introduction of the act of 85/100 (95% CI: 82 to 89) (George, 2006; Hasketh & Xing, 2006; Sahni et al, 2008; Egan et al, 2011).

Natural history

At the time a diagnosis of DMD is made, young boys with the disease have few complaints apart from an inability to keep up with their peers and occasional muscle tenderness and calf stiffness following exercise. As the disease progresses muscle weakness becomes more profound and musculotendinous contractures of the flexor muscles of the elbows, knees, hamstrings and hips develop, which limits the range of motion in the associated joints. These contractures occur due to imbalances in muscle strength between the flexor and extensor muscles, with the stronger muscles on one side of the joint stretching their weaker antagonists. A wheelchair is initially only required for long distances outside of the home but as the disease progresses independent ambulation becomes increasingly more difficult, and eventually wheelchair use becomes permanent. In a case series of 120 boys reported by Emery & Muntoni (2003) 10% became wheelchair bound by 6.7 years of age, 50% by 8.5 years and 99% by 13.2 years of age. Once ambulation is lost, up to 90% of individuals develop a progressive scoliosis with pelvic obliquity and a loss of sitting balance, as a consequence of the collapse of the spinal musculature (Alman & Kim, 1999; Heller et al, 2001; Biggar et al, 2002; Manzur & Muntoni, 2002; Sengupta et al, 2002; Emery & Muntoni, 2003).

The onset of cardiac involvement in DMD typically occurs between 14 to 16 years of age, although in most it remains subclinical until the late stages of the disease. The first manifestations of cardiac involvement present as sinus tachycardia, ECG abnormalities and arrhythmias. As the disease progresses the cells of the myocardium are replaced by fibrous and adipose tissue and there is increasing evidence from echocardiography of myocardial thickening, dilation of the cardiac cavities, valve abnormalities and abnormal muscle wall motion. In approximately 10-20% of individuals cardiac involvement is so severe that it results in death at approximately 16 years of age. Smooth muscle involvement in the disease is not well studied, although dysphasia, severe constipation, bladder dysfunction, paralytic ileus, abdominal pain and distension, acute gastric dilatation and pseudo-obstruction have been clearly documented in case histories. This is supported by evidence obtained from detailed autopsies of the gastrointestinal tract of individuals with DMD, that have demonstrated the presence of a pattern of smooth muscle fibre damage.
like that seen in skeletal and cardiac muscle, with muscle fibre atrophy, variations in fibre size and fibrosis, and a loss of muscle fibres particularly in the distal esophagus, stomach, small bowel and colon (Quinlivan et al, 1996; Fois, 1997; Hoffman, 2001; Manzur & Muntoni, 2002; Emery & Muntoni, 2003; Muntoni, 2003; Finsterer & Stollberger, 2003).

In 1868 Duchenne described five patients with DMD who he believed had some degree of cognitive impairment. It is now known that the mean intelligence quotient of individuals with DMD is approximately 82, 1 standard deviation below the population mean of 100, with 30% of cases having intelligence quotients below 75. Irrespective of their general intelligence level, individuals with the disease perform poorly on tests of verbal comprehension and story memory skills. During early childhood respiratory function and respiratory muscle strength in individuals with DMD increases relatively normally parallel to the growth of the thorax and lungs. As the disease progresses there is an early loss of expiratory muscle strength, followed by a loss of inspiratory muscle strength. These combined losses ultimately counteract the effects of the growth of the thorax and lungs and respiratory function reaches a plateau at the time ambulation is lost. Within a year of reaching this plateau respiratory function falls to 50% below that predicted for height, age and sex. It continues to decline from this point until respiratory function falls below 20% of predicted and chronic respiratory failure develops, and unless some form of LTMV is used.

---

5 During normal inspiration, what is known as tidal or quiet supine breathing, the respiratory muscles increase the volume of the thorax by changing its diameter. This reduces the pressure of the air in the lungs relative to that of the surrounding atmosphere, which results in the flow of air into the lungs to equalize the pressure differential. Inspiration ends once the pressure of air in the lungs returns to that of the surrounding atmosphere. The diaphragm, a dome like muscle that sits between the abdomen and thorax, generates between 60 to 70% of the negative intrathoracic pressure required for (tidal) breathing. The pressure generated by the diaphragm is augmented by the external intercostals which lift the rib cage, swinging it outwards and pulling the sternum forward, expanding the diameter of the thorax laterally and anteroposteriorly. As respiratory demand increases the accessory inspiratory muscles, a group of muscles comprised of the external intercostals, the parasternal sections of the internal intercostals, the triangularis sterni and the scalenes, are recruited to further expand the thorax by maximizing rib cage expansion. Under conditions of severe respiratory demand the muscles of the neck, upper back and shoulders, specifically the sternomastoid, subclavian, pectoralis minor and major, serratus anterior, upper and lower trapezius and latissimus dorsi muscles, are brought into play (Marieb, 1998; Gibson, 1999; Polkey & Moxham, 1999; Aldrich & Rochester, 2000; Celli, 2001). During normal (unforced) expiration the inspiratory respiratory muscles relax, and the elastic recoil of the lungs and chest wall return the thorax to its resting volume. The decrease in volume in the thorax increases the pressure of air within the lungs to above that of the surrounding atmosphere, which results in air flowing out the lungs to equalize the pressure differential. When sitting and standing, the scalene and parasternal muscles add to the elastic recoil of the lungs and chest wall, and low level abdominal muscle activity begins to act to assist expiratory airflow. The pressure required for active (forced) expiration, for activities such as coughing, is principally generated by internal oblique and the transversus abdominis muscles. The contraction of these muscles raises intra-abdominal pressure forcing the abdominal organs against the diaphragm. The rectus abdominis acts to restrict the movement of these organs, with the lateral internal intercostals and transversus thoracis (triangularis sterni) further increasing expiratory pressure. The quadratus lumborum, the internal intercostals and the latissimus dorsi depress the rib cage, further decreasing the size of the thorax and increasing the pressure of the air within the lungs (Marieb, 1998; Gibson, 1999; Polkey & Moxham, 1999; Aldrich & Rochester, 2000; Celli, 2001; Bach, 2002b).

6 In DMD there are two primary causes of death: (i) cardiac, the result of clinically evident severe cardiomyopathy at approximately 16 years of age and (ii); respiratory, the result of chronic respiratory failure secondary to severe respiratory muscle weakness at approximately 19 years of age (without ventilation). That is not to say that
this will result in death in the late teens or early twenties (Hahn et al, 1997; Simonds et al, 1998; Hinton et al, 2000; Felisari et al, 2000; Schramm, 2000; Cotton, Voudouris & Greenwood, 2001; Anderson et al, 2002; Emery, 2002b; Biggar et al, 2002; Emery & Muntoni, 2003; American Thoracic Society, 2004; Cyrulnik et al, 2008).

Hahn et al (1997) undertook a study of the decline of respiratory function and respiratory muscle strength of 52 patients with DMD. The diagnosis of the disease was made using the standard diagnostic criteria in use at that time, which did not include DNA analysis. It was acknowledged by the authors that this meant that it was possible that severe cases of Becker muscular dystrophy had been mistakenly included in the study. The study population was made up of consecutive patients; 17 (32.69 %) who were 7 to 14 years of age, 15 (28.9 %) who were 14 to 18 years of age, and 20 (38.5 %) who were older than 20 years of age. There were 23 individuals (44.2 %) with maximal inspiratory pressures below 40% of predicted for height, age and sex. Long term mechanical ventilation was used to support respiratory function in 22 (95.7 %) of the individuals with maximal inspiratory pressure below 40% of predicted. The absolute vital capacity of the study participants was to shown to ascend up to the point ambulation was lost, reaching a plateau at a mean age of 14 years from where it declined at a mean rate of 7.9% per year.

Table 3 lists in detail the decline in vital capacity in relation to age reported in the study. In this study maximal inspiratory pressure (MIPs) and maximal expiratory pressure (MEPs) were found to increase from 7 to 14 years of age, then decrease sharply from 15 to 18 years after which the rate of decline slowed, as shown in Table 4.

individuals who die from severe cardiomyopathy do not have some degree of respiratory impairment but simply that it was not severe enough in itself to be the primary cause of death.

There are four primary measures of static lung volume: (i) tidal volume; (ii) inspiratory volume; (iii) expiratory reserve volume; and (iv) residual volume and four primary measures of lung capacity: (i) total lung capacity; (ii) vital capacity; (iii) inspiratory capacity; and (iv) functional residual capacity. These measures, which are defined in Appendix 1, have population standard values, that vary with the height, age and sex of an individual (Marieb, 1998; Polkey & Moxham, 1999; Aldrich & Rochester, 2000; American Thoracic Society, 2004; Pierce et al, 2005).

The study was conducted prior to research that demonstrated that severe forms of Limb Girdle Type 21 muscular dystrophy could also be misclassified as DMD.

The combined (global) strength of the respiratory muscles is measured indirectly as the pressure of air generated at the mouth. The two most commonly used measures are: maximal static inspiratory pressure (MIPs); and maximal static expiratory pressure (MEPs). In normal healthy young males maximal inspiratory pressure is greater than 120 cm H₂O, and maximal expiratory pressure is greater than 150 cm H₂O (Marieb, 1998; Polkey & Moxham, 1999; Aldrich & Rochester, 2000; American Thoracic Society, 2004; Pierce, Hillman, Young, O'Donoghue, Zimmerman, West & Burdon, 2005).
### Table 3: Vital capacity and total lung capacity in relation to age

<table>
<thead>
<tr>
<th>Age (yrs)</th>
<th>Absolute vital capacity</th>
<th>Absolute vital capacity %&lt;sup&gt;10&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 – 8</td>
<td>1390 mls (297)</td>
<td>92 % (28)</td>
</tr>
<tr>
<td>9 – 10</td>
<td>1851 mls (434)</td>
<td>84 % (22)</td>
</tr>
<tr>
<td>11 – 12</td>
<td>2223 mls (471)</td>
<td>88 % (21)</td>
</tr>
<tr>
<td>13 – 14</td>
<td>2658 mls (870)</td>
<td>76 % (14)</td>
</tr>
<tr>
<td>15 – 16</td>
<td>1836 mls (994)</td>
<td>49 % (21)</td>
</tr>
<tr>
<td>17 – 18</td>
<td>1374 mls (858)</td>
<td>31 % (19)</td>
</tr>
<tr>
<td>19 – 20</td>
<td>1199 mls (944)</td>
<td>25 % (18)</td>
</tr>
<tr>
<td>21 – 22</td>
<td>808 mls (496)</td>
<td>18 % (11)</td>
</tr>
<tr>
<td>23 yrs +</td>
<td>459 mls (408)</td>
<td>14 % (7)</td>
</tr>
</tbody>
</table>

Source: Hahn et al (1997). All values expressed as mean millilitres (standard deviation)

### Table 4: Maximal respiratory pressures in relation to age

<table>
<thead>
<tr>
<th>Age (yrs)</th>
<th>MIPs max</th>
<th>MIPs max %&lt;sup&gt;11&lt;/sup&gt;</th>
<th>MEPs max</th>
<th>MEPs max %</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 – 8 years</td>
<td>35 (7)</td>
<td>56 (14)</td>
<td>39 (9)</td>
<td>53 (12)</td>
</tr>
<tr>
<td>9 – 10 years</td>
<td>42 (9)</td>
<td>55 (18)</td>
<td>39 (9)</td>
<td>44 (10)</td>
</tr>
<tr>
<td>11 – 12 years</td>
<td>56 (3)</td>
<td>77 (4)</td>
<td>47 (13)</td>
<td>48 (13)</td>
</tr>
<tr>
<td>13 – 14 years</td>
<td>64 (8)</td>
<td>77 (18)</td>
<td>56 (5)</td>
<td>52 (5)</td>
</tr>
<tr>
<td>15 – 16 years</td>
<td>43 (14)</td>
<td>52 (16)</td>
<td>32 (11)</td>
<td>27 (9)</td>
</tr>
<tr>
<td>17 – 18 years</td>
<td>32 (16)</td>
<td>30 (17)</td>
<td>26 (12)</td>
<td>18 (10)</td>
</tr>
<tr>
<td>19 – 20 years</td>
<td>33 (22)</td>
<td>28 (18)</td>
<td>28 (18)</td>
<td>18 (12)</td>
</tr>
<tr>
<td>21 – 22 years</td>
<td>26 (17)</td>
<td>22 (14)</td>
<td>19 (10)</td>
<td>12 (6)</td>
</tr>
<tr>
<td>&gt;=22 years</td>
<td>21 (10)</td>
<td>18 (8)</td>
<td>23 (20)</td>
<td>15 (12)</td>
</tr>
</tbody>
</table>

Source: Hahn et al (1997). All values expressed as mean cm H₂O (standard deviation)

---

<sup>10</sup> As a percentage of the population norm, predicted for height, age and sex

<sup>11</sup> Percentage of predicted for height, age and sex
The loss of ambulation in individuals with DMD, prior to the point at which respiratory function starts to decline, acts to reduce the demands placed on their respiratory systems and consequently at this time, there are few respiratory complaints and individuals do not feel short of breath. The first signs of the decline of respiratory function are instead seen during sleep which is disrupted by abnormal breathing, what is known as sleep-disordered breathing (SDB). The onset of SDB occurs many years before the onset of clinically apparent daytime respiratory insufficiency, a result of the combination of inspiratory muscle weakness, the general reduction in muscle tone that occurs during sleep and the loss of the “wakefulness drive to breathe”. This “wakefulness drive to breathe” is the ability, while awake, to partially compensate for respiratory muscle weakness through the recruitment of the accessory respiratory muscles (Sivak, Shefner & Sexton, 1999; Attarian, 2000; Langevin et al, 2000; Chokroverty, 2001; Bourke & Gibson, 2002; Piper, 2002; Wagner, Lechtzin & Judge, 2007).

In humans there are two main types of sleep, non-rapid eye movement (NREM) and rapid eye movement (REM) sleep. During a normal night’s sleep, for a healthy young male, there are approximately 4 to 5 ninety minute cycles of alternate NREM and REM sleep, with the proportion of each cycle spent in REM increasing as the night progresses. During each of these sleep cycles there is a reduction in muscle tone, temperature regulation, blood pressure, heart rate and central nervous system activity, with the greatest reductions seen during REM sleep. In NREM sleep there is a slight reduction in tone of the respiratory muscles and the response of the respiratory system to arterial carbon dioxide levels. In REM sleep there is a general reduction in muscle tone that spares only the intra-ocular muscles and the diaphragm, and a significant reduction in the response of the respiratory system to arterial carbon dioxide levels (Levy et al, 1998; Langevin et al, 2000; Chokroverty, 2001; Bourke & Gibson, 2002; Piper, 2002).

Sleep-disordered breathing in individuals with DMD initially only occurs during REM sleep, when the load placed on the diaphragm exceeds its capacity to maintain adequate ventilation. This mismatch between load and capacity results in apnoeas and hypopnoea (definitions in Table 5), events that are accompanied by a decrease in oxygen saturation and an increase in carbon dioxide levels. These events may be terminated by arousals, non-specific protective responses to a range of stimuli during sleep that re-enable the use of the accessory respiratory muscles to return arterial blood gases to acceptable levels. While these arousals reduce the severity of the drop in oxygen saturation and the increase in carbon dioxide levels, they have a detrimental effect on the duration and quality of sleep, which can result in daytime symptoms such as fatigue, excessive sleepiness and cognitive impairment. Over time increasing pressure to achieve and maintain sleep, particularly REM sleep, results in a blunting of the sensitivity of the respiratory centre to increasing carbon dioxide levels and there is a consequent reduction in the number of these arousals with progressively longer periods of severe oxygen desaturation and abnormally high
carbon dioxide levels during sleep. If left untreated, this blunting of the response of the respiratory centre to carbon dioxide levels will progress to diurnal hypercapnia\textsuperscript{12}, respiratory failure and death (Levy et al, 1998; Kirk et al, 2000; Langevin et al, 2000; Chokroverty, 2001; Zucconi & Bruni, 2001; Piper, 2002).

**Table 5: Abnormal respiratory events during sleep\textsuperscript{13}**

<table>
<thead>
<tr>
<th>Event</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apnoea</td>
<td>Absence of airflow for more than 10 secs</td>
</tr>
<tr>
<td>Hypopnoea</td>
<td>Reduction in airflow or thoracoabdominal effort for &gt;10 seconds accompanied by &gt; 3 % oxyhemoglobin desaturation or EEG arousal of &gt; 3 seconds</td>
</tr>
<tr>
<td>Apnoea / hypopnoea index (AHI)\textsuperscript{14}</td>
<td>Number of apnoeas and hypopnoea divided by total sleep time\textsuperscript{15}</td>
</tr>
<tr>
<td>Hypoventilation</td>
<td>Oxygen desaturations that are associated with the transient elevation of transcutaneous carbon dioxide levels of &gt; 10 mm HG</td>
</tr>
</tbody>
</table>


Phillips et al (2001) undertook a retrospective study of the cause of death in 58 individuals with DMD. The main aim of the study was to determine whether the rate of decline in vital capacity could be used as a predictor of survival. The vital capacity of study participants was monitored from 12 years of age (median, range 10 to 18) for a period of 7.2 years (median, range 2 to 16.3 years). At baseline vital capacity was 1.6 litres (median, n=58, range 0.45 to 2.75). The vital capacity of the study participants was measured a total of 523 times during the study period (median of 8.5 measurements per individual, range 2 to 16). At the age of 14 vital capacity was 1.8 litres (median, n=34, range 0.55 to 3.5) or 50% of that predicted for height, age and sex. At 18 years of age vital capacity was 1.4 litres (median, n=31, range 0.4 to 3.7) or 36% of predicted. A total of 37 (63.79 \%) study participants died during the data collection period at a median age of 20.5 years, with 4 (10.81 \%) of the deaths due to cardiac causes. The only variable that was shown to be predictive of survival was the age at which vital capacity declined to 1000 mls (Cox

\textsuperscript{12} Arterial carbon dioxide level of greater than 45mmHg (Fanfulla et al, 1997)

\textsuperscript{13} These definitions can vary significantly between studies, with consequent variability in the events recorded by studies and the conclusions drawn (Tang et al, 2002).

\textsuperscript{14} Also known as the respiratory disturbance index or RDI (Tang et al, 2002).

\textsuperscript{15} Or a specified stage of sleep, e.g. during REM sleep.
regression analysis, \( r=0.94, n=37, p <0.0003 \). The median period of survival once this point was reached was 3.1 years with only 8 % of patients surviving for 5 years.

**Clinical management**

The clinical management of DMD has as its aim the early detection and treatment of complications. This requires a multidisciplinary team approach with one physician co-ordinating the provision of care provided by a team ideally made up of pediatricians, neurologists, orthopaedic surgeons, cardiologists, respiratory specialists, physiotherapists, orthotists, occupational therapists, speech and language therapists, dieticians, psychologists and social workers (Manzur & Muntoni, 2002; Bushby et al, 2010a).

In the early ambulant stages of the disease the mainstay of treatment is physiotherapy, with passive stretching, static positioning of limbs, and the use of individualised orthoses to prevent or reduce the severity of contractures. While corticosteroids are often used alongside physiotherapy to maintain muscle strength and to delay the loss of ambulation and the development of scoliosis, there is no consensus on their use as a standard treatment. The progression of muscle weakness after the loss of ambulation eventually necessitates the need for surgical tendon release to correct knee and hip contractures, and to correct scoliosis, with the aim of improving comfort and functional status. Scoliosis is corrected through the use of an internal fixation system of stainless steel rods that are inserted parallel to the spine and individually wired to each vertebra. The rods extend from the upper thoracic spine, at the level of T2-T4 to either the lumbar vertebra (L5) or to the sacrum or pelvis. While there is evidence from an early age of cardiac abnormalities in all individuals with DMD, it is not until left ventricular function falls to below 10-15% of normal, and there is overt cardiac failure, that this becomes clinically apparent. For this reason, cardiac function is evaluated once every six months from the age of 10 in all individuals with DMD. Prior to the development of overt cardiac failure, angiotensin converting enzyme inhibitors and low doses of non-selective beta-adrenergic blocking drugs are used to improve left ventricular function. When left ventricular function falls to below 10-15% of normal, cardiac glycosides such as Digitoxin are used to improve cardiac output, with loop diuretics such as Frusemide and Spironolactone used to provide symptomatic relief (Bakker & Van Ommen, 1998; Alman & Kim, 1999; Bakker et al, 2000; Emery, 2002a; Emery, 2002b; Manzur & Muntoni, 2002; Meola, Kaparti & Griggs, 2002; Thacker et al, 2002; Emery & Muntoni, 2003; Finsterer & Stollberger, 2003; Manzur et al, 2004; Wagner, Lectzin, & Judge, 2007; Bushby et al, 2010a; Bushby et al, 2010b).
Prior to the loss of ambulation the nutritional status of individuals is assessed every six months through the measurement of height and weight. Once ambulation is lost, and individuals are no longer able to stand, upper arm length, tibial length, or knee height is used as proxy measures for height. The levels of vitamins and other dietary micronutrients are monitored, and Vitamin D, calcium and other supplements are prescribed as required. As the disease progresses the development of pharyngeal muscle weakness can result in dysphagia and episodes of choking which can lead to malnutrition and dehydration. To prevent this from happening, families are given advice about how to improve or maintain the oral intake of foods and liquids through the use of simple measures, such as seating individuals bolt upright during meals, or when drinking, to reduce their risk of aspiration or choking. Food can also be cut into small pieces, or texture modified, to make it easier to eat and energy intake can be increased through the use of caloric dietary supplements. If there is evidence of an unintentional weight loss of greater than 10%, fever of unknown cause, persistent coughing, gagging or choking during meals, or a history of aspiration pneumonia, then individuals are sent for a videofluoroscopic swallowing study. In a minority of cases swallowing may be so severely impaired that a gastrostomy feeding tube is needed to maintain adequate nutritional status (Manzur & Muntoni, 2002; American Thoracic Society, 2004; Davidson & Truby, 2009; Bushby et al, 2010a).

In non-ambulant individuals with DMD there is also a risk of the development of obesity which can be very difficult to manage, as the standard approach of a reduction in energy intake and an increase in physical activity levels cannot be easily undertaken and is not without risk, as it may result in an increase in the rate in which muscle mass is lost. The nutritional management of DMD is further complicated by the progressive decline in lean body mass that is the result of the replacement of muscle tissues with adipose and connective tissue, which makes it very difficult to assess an individual's nutritional status and their energy requirements. There is also very little high-quality evidence, or consensus guidelines, that dieticians can use to guide their management of individuals with DMD. Furthermore, in clinical practice, the management of the nutritional status of individuals with DMD is overwhelmed by the need to address the decline in respiratory function, even though malnutrition and obesity are in themselves significant health issues that are known to have an negative impact on respiratory muscle strength and function (Manzur & Muntoni, 2002; American Thoracic Society, 2004; Davidson & Truby, 2009; Bushby et al, 2010a).

Respiratory failure is the main cause of death in DMD, and consequently the early detection and treatment of respiratory complications is an important component of its clinical management. Routine respiratory assessments and testing of respiratory function and muscle strength is undertaken at least annually from 8 years of age. Once ambulation is lost, the frequency of routine respiratory assessment is increased to at
least six monthly, with the addition of the measurement peak cough flow\textsuperscript{16} to assess the individual’s ability to produce an effective cough. An effective cough is an essential component of the respiratory system, one that protects the lungs against aspiration and removes excessive secretions from the airways. The combination of an ineffective cough and declining respiratory function in an individual with DMD can mean that even a relatively minor upper respiratory tract infection has the potential to rapidly progress to severe and potentially life threatening pneumonia, respiratory muscle fatigue, acute respiratory failure and death. In DMD the ability to produce an effective cough is lost relatively late in the disease when inspiratory muscle weakness severely reduces the volume of air that can be inspired. For as long as a good initial inspiratory effort can be made, the elastic recoil of the lungs and chest wall is enough to ensure that an effective cough can be produced (Bach, Ishikawa & Kim, 1997; Binkrant, Pope & Eiben, 1999; Leger & Hill, 2001; American Thoracic Society, 2004; Perrin et al, 2004; Wagner, Lechtzin & Judge, 2007; Bushby et al, 2010b).

Annual influenza vaccination and manually assisted coughing techniques are used to reduce the risk of pneumonia, and to clear respiratory secretions during established infections. Manual assisted coughing is performed with the assistance of physiotherapists, family members and other carers, with a self-inflating resuscitator or glossopharyngeal breathing used to increase inspiratory capacity, and abdominal thrusts or lateral rib cage compression used to augment the pressure generated to produce an effective cough. Antibiotics are used to aggressively to treat established infections, without the need for evidence from sputum cultures, if oxygen saturation is less than 95% on room air. Non-invasive ventilation may also be needed to support ventilation during an chest infection, with oxygen therapy used with caution, as it may worsen hypercapnia and impair central respiratory drive in individuals with pre-existing long term respiratory failure (Benditt & Boitano, 2005; Wagner, Lechtzin & Judge, 2007; Bushby et al, 2010b).

The diagnosis and staging of sleep-disordered breathing requires an admission to a specialist sleep study unit for an overnight examination (polysomnographic evaluation). In a sleep study the presence of sleep-disordered breathing, and its severity, is established through the continuous monitoring of a number of physiological variables that determine the sleep state and stage, the respiratory effort being made and the effectiveness of gas exchange. The data collected during the sleep study is summarized into a set of numbers of abnormal respiratory events, such as the number of periods of hypoventilation, apnoeas, hypopnoeas per hour of sleep (as defined in Table 5) (Attarian, 2000; Chokrovery, 2001; Whiteford, Fleming & Henderson, 2004).

\textsuperscript{16} The maximum expiratory flow of gas produced during a cough manoeuvre (Suarez et al, 2002)
The onset of sleep-disordered breathing in DMD is insidious and its early signs and symptoms: nocturnal restlessness; frequent unexplained arousals from sleep; daytime fatigue or sleepiness; morning headaches; impaired concentration; and emotional and behavioural disorders, are subtle and non-specific. Individuals with DMD (and their parents) are routinely asked by their doctors about the quality of sleep and the presence of signs and symptoms that are suggestive of sleep-disordered breathing. Although the subtle and non-specific nature of these symptoms, and their denial by patients or family members who perceive them to be pre-terminal events, means that sleep-disordered breathing may be present even if there are no observed or reported signs or symptoms. Ideally, sleep studies should be conducted annually from the point at which ambulation is lost; with awake arterial blood gas analysis performed six monthly to measure oxygen saturation and carbon dioxide levels. Sleep studies are a relatively expensive form of investigation that is inconvenient for patients and their families and consequently before ordering a study clinicians need to be fairly certain that an individual does in fact have the problem (Labanowski, Schmidt-Nowara & Guileminault, 1996; Hukins & Hillman, 2000; Piper, 2002; American Thoracic Society, 2004; Benditt & Boitano, 2005; Wagner, Lecthizin & Judge, 2007).

A number of studies, such as those of Barbe et al (1994), Hukins & Hillman (2000) and Kirk et al (2000) have been conducted since the early 1990s with the aim of developing objective measures of daytime respiratory function and/or respiratory muscle strength that could be used by clinicians to predict the presence and severity of sleep-disordered breathing. The only study to date that has been able to do this is that of Mellies et al (2003). In this study there were 49 children and adolescents with a range of neuromuscular diseases, 22 girls (44.9%) and 27 boys (55.1 %), with a mean age of 11.3 years (SD 4.4). There were 7 individuals (14.3 %) with DMD with a mean age of 14.6 years (SD 4.0). At the time of their entry into the study, none of the study participants was using LTMV or had been previously been diagnosed with sleep-disordered breathing. From the finding of the sleep studies of these 49 individuals the authors reported that: (i) the onset of sleep-disordered breathing in REM, without hypercapnia, could be predicted from a vital capacity of less than 60 % of predicted or a maximal inspiratory pressure of less than 30 mmHg (4 kPa) and; (ii) the onset of sleep-disordered breathing in REM with hypercapnia could be predicted from a vital capacity of less than 40 % of predicted, or a maximal inspiratory pressure of less than 19 mmHG (2.5 kPa).

The use of long-term mechanical ventilation\textsuperscript{17} to support respiratory function in individuals with DMD is undertaken to achieve both short-term and long-term goals. The short-term goals are the relief of the symptoms of sleep-disordered breathing and breathlessness, and to improve arterial blood gases. The

\textsuperscript{17} The use of mechanical ventilation for more than 6 hours a day, for a period of greater than 3 months (Make, 2001; Lloyd-Owen et al, 2005).
long-term goals are the maintenance of arterial blood gases to within acceptable limits, the relief of the symptoms of sleep-disordered breathing and breathlessness, the maintenance of improvement of quality of life, and most importantly the prolongation of survival. Unfortunately there are no clear and consistent guidelines for the use of LTMV in DMD, which makes it very difficult for clinicians to know when to initiate the intervention. European consensus guidelines recommend initiating LTMV when daytime arterial carbon dioxide levels rise above 45 mmHg. American consensus guidelines only suggest that clinicians consider initiating treatment if: (i) there are daytime symptoms of sleep-disordered breathing; (ii) daytime arterial carbon dioxide levels are above 45 mmHg; (iii) nocturnal oxygen saturation, measured by a pulse oximeter, falls to 88% or less for five consecutive minutes during sleep; (iv) maximal inspiratory pressure is less than 60 cm H2O; (v) or vital capacity is less than 50% of predicted. In Australia, guidelines developed for use in New South Wales (Wark, Murray & Flunt, 2010) include the following as indications for a trial of domiciliary LTMV, if: (i) there are symptoms of significant sleep-disordered breathing; (ii) awake carbon dioxide levels are greater than or equal to 45 mmHg; (iii) there is a sustained fall in arterial oxygen levels during sleep to below 90% for more than 5% of total sleep time; (iv) there is an increase in arterial carbon dioxide levels during sleep of 8 mmHg or higher; (v) there is a peak carbon dioxide level of greater than 50 mmHg for more than 50% of total sleep time; (vi) a vital capacity of less than 50% of predicted or; (vii) a maximal inspiratory pressure of less than 40% of predicted (Hammer, 2000; Metha & Hill, 2001; American Thoracic Society, 2004; Langmack & Make, 2002; Fiorenza, Vitacca & Clini, 2003; Benditt & Boitano, 2005; Toussaint, Chatwin & Soundon, 2007).

The provision of ongoing psychosocial support for individuals with DMD, and their parents and siblings, is an important component of the management of the disease. For while many individuals with DMD may give the impression of having come to terms with their disease, the inevitability of death at a young age places a significant emotional and psychological burden on them. Family members, typically mothers, who are called upon to increasingly provide specialised and time consuming care for their disabled sons, must do this while experiencing a range of emotions, such as separation, loss, sadness, guilt and anger. The wellbeing of non-affected siblings can also be impacted upon, as their parent’s focus on their affected sibling leaving them alone to deal with feelings of devastation, jealousy and guilt. Individuals with DMD and their families rarely discuss the terminal nature of the disease or put in place any advance end-of-life directives. Families are generally reluctant to discuss issues related to end-of-life care with the healthcare team, and may simply refuse to discuss it at all. It is very unusual for children to die, and consequently services for the provision of end-of-life care for children are often fragmented and ad-hoc, and are generally less well organised than services for adults. In the terminal stages of DMD upper respiratory tract infections can rapidly progress to severe pneumonia and death, and consequently home palliative care, or hospice care may be not required. Unfortunately, without advance end-of-life directives and a
formal referral to palliative care services, families and carers may not be able to easily gain access to professional bereavement services after the death of a son with DMD (Polakoff et al, 1998; Parker, Maddocks & Stern, 1999; Reid & Reinwich, 2001; Hedderly, Baird & McConachie, 2002; Issacs & Sewell, 2003; Nereo, Fee & Hinton, 2003; American Thoracic Society, 2004; Simonds, 2004; Erby, Rushton & Geller, 2006; Bushby et al, 2010a; Penner, Cantor & Seigel, 2010).

The next table, Table 6, summarises the ongoing clinical management of DMD after ambulation is lost, prior to and after the initiation of LTMV.

**Table 6: Clinical management from the loss of ambulation, prior to & after the initiation of LTMV**

<table>
<thead>
<tr>
<th>Team member</th>
<th>Activity</th>
<th>Pre initiation of LTMV</th>
<th>Post initiation of LTMV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurologist</td>
<td>Team co-ordination</td>
<td>Six monthly</td>
<td>Six monthly</td>
</tr>
<tr>
<td></td>
<td>General monitoring</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiologist</td>
<td>Monitoring of cardiac function</td>
<td>Six monthly</td>
<td>Six monthly</td>
</tr>
<tr>
<td></td>
<td>Diagnosis and treatment of abnormalities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dietician</td>
<td>Monitoring of nutritional status</td>
<td>Six monthly</td>
<td>Six monthly</td>
</tr>
<tr>
<td></td>
<td>Dietary supplementation as required</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social worker / Psychologist</td>
<td>Provision of psychosocial support for the patient &amp; their family</td>
<td>Six monthly</td>
<td>Six monthly</td>
</tr>
<tr>
<td>Respiratory specialist</td>
<td>Monitoring of the decline of respiratory function &amp; respiratory muscle strength</td>
<td>Six monthly</td>
<td>Six monthly</td>
</tr>
<tr>
<td></td>
<td>Measurement of peak cough flow</td>
<td>Six monthly</td>
<td>Six monthly</td>
</tr>
<tr>
<td></td>
<td>Arterial blood gas analysis</td>
<td>Annual</td>
<td>Six monthly</td>
</tr>
<tr>
<td></td>
<td>Sleep studies</td>
<td>Annual diagnostic study</td>
<td>Six monthly titration study</td>
</tr>
</tbody>
</table>
Alternative treatment pathways

There are four distinct stages in the evolution of respiratory failure in DMD. In the first stage breathing is disordered during REM sleep but carbon dioxide levels remain within normal limits (Sleep-disordered breathing in REM sleep without hypercapnia). In the second stage sleep is disordered during REM sleep and carbon dioxide levels rise above normal limits (Sleep-disordered breathing in REM, with hypercapnia). In the third stage breathing is disordered during both REM and NREM sleep, and carbon dioxide levels remain above normal limits for the entire sleep cycle but return to within normal limits during the day (Sleep-disordered breathing in REM and Non-REM sleep with hypercapnia). In the fourth and final stage, carbon dioxide levels are above normal limits during the day, and further deteriorate during sleep (Diurnal hypercapnia) (Simonds et al, 1998; Ward & Hill, 2001; Zucconi & Bruni, 2001; Emery & Muntoni, 2003; American Thoracic Society, 2004; Toussaint, Chatwin & Soudon, 2007).

There are three distinct alternative treatment pathways for the timing of the initiation of LTMV to support respiratory function in individuals with DMD:

(i) To not use it at all;
(ii) To initiate the intervention from the onset of SDB in REM, without hypercapnia and;
(iii) To initiate the intervention from the onset of SDB in REM, with hypercapnia (Toussaint, Chatwin & Soudon, 2007).

The alternative treatment pathways outlined above differ from the current practice use of LTMV in DMD only in relation to the timing of initiation and the duration of treatment. All other aspects of the current practice clinical management of DMD are unchanged. The first of the three options, that of not using it all is not recommended practice in countries where the technology is available. The evidence obtained from observational studies conducted over the last 25 years for the use of LTMV for neuromuscular diseases and chest wall disorders, is such that to not offer to provide the intervention as a possible treatment option may be seen as a breach of medical ethics. The technology is available in Australia, and consequently this alternative treatment pathway was not examined in the economic evaluation (Turkington & Elliot, 2000; Shneerson & Simonds, 2002; Toussaint, Chatwin & Soudon, 2007).
In the thesis the incremental Cost-effectiveness and cost-utility of alternative treatment pathways two and three were examined in an economic evaluation. In which the: (i) initiation of LTMV from the onset of SBD in REM without hypercapnia (Intervention 1); and (ii) the initiation of LTMV from the onset of SDB in REM with hypercapnia (Intervention 2) were compared to current practice initiation of LTMV from the onset of diurnal hypercapnia. The initiation of LTMV in DMD prior to the onset of diurnal hypercapnia was, until recently, not generally well accepted. For while it has been claimed by authors such as Rideau, Delaubier, Guillou & Renarde-Irani (1995) that the early initiation of LTMV could slow the rate of respiratory muscle decline, and consequently prolong survival, there is very little evidence to support this. Furthermore, evidence from the only randomised controlled trial of the use of LTMV in DMD that of Raphael et al (1994), reported that the early initiation of the intervention decreased life expectancy. While the findings of the study of Raphael et al (1994) have been widely questioned due to a range of methodological weaknesses, for many years after its publication there was little exploration of the early initiation of LTMV in DMD (Leger & Hill, 2001; Eagle et al, 2002; Simonds, 2002; Toussaint, Chatwin & Soudon, 2007).

A renewed interest in the initiation of LTMV in DMD prior to the onset of diurnal hypercapnia is part of an increasingly pro-active approach to the treatment of the two predictable life threatening complications of the disease: (i) severe cardiomyopathy and; (ii) respiratory failure. It is possible that the early initiation of LTMV, such as from the onset of SBD in REM without hypercapnia or from the onset of SDB in REM with hypercapnia, may be able to maintain or improve the quality of life of individuals with DMD. As there are currently no available interventions or therapies that have any significant effect on the underlying disease process in DMD, interventions such as LTMV that may be able to maintain, or improve, quality of life (while prolonging survival) are of value as they can be of benefit to patients and their families (Eagle et al, 2002; Toussaint, Chatwin & Soudon, 2007; Muntoni, 2010).

As part of the renewed interest in the early initiation of LTMV in DMD, Ward, Chatwin, Heather and Simonds (2005) conducted a randomised controlled study of the early initiation of LTMV for individuals with neuromuscular diseases and chest wall disorders. The aim of the study was to test whether the early initiation of LTMV was a valid treatment option. The primary study outcome was daytime peak arterial carbon dioxide levels, which were measured six monthly for a 24 month period. There were 12 (27.08%) individuals with DMD in the study population of 48. At baseline all of the study participants had vital capacities of less than 50% of predicted and symptoms suggestive of sleep disordered breathing. Baseline sleep studies found that twenty-six (54.17 %) of the study participants had SBD with hypercapnia in REM, and they were randomized to receive usual care, or usual care with LTMV. Nineteen of the study participants (39.58%) were found to have diurnal hypercapnia and they were immediately treated with
LTMV. There were 3 individuals with DMD in the control group, 2 in the experimental group, and 8 in the immediate treatment group. After 24 months there were no deaths in any of the study groups, and no statistically significant differences in the daytime peak arterial carbon dioxide levels of the control and experimental groups.
Conclusion

Duchenne muscular dystrophy is a currently incurable X-linked disorder that is invariably fatal at a relatively young age. The main cause of death in the disease is respiratory failure secondary to severe respiratory muscle weakness. The decline of respiratory function in the disease begins within a year of the loss of ambulation, with the first manifestations of this decline seen during sleep in the form of sleep-disordered breathing. As respiratory function worsens, so does the severity of sleep-disordered breathing, and there are increasingly frequent arousals from sleep with resultant daytime fatigue, excessive sleepiness, cognitive impairment, and a reduction in quality of life.

In recent years there has been renewed interest in the early initiation of long-term mechanical ventilation in DMD as part of an increasingly pro-active approach to the treatment of the disease. In this thesis, the Cost-effectiveness and cost-utility of the initiation of LTMV from the onset of SBD in REM without hypercapnia (Intervention 1) and, the initiation of LTMV from the onset of SDB in REM with hypercapnia (Intervention 2) were compared to the initiation of LTMV in current practice.

In the next chapter of the thesis, there is an introduction to economic evaluation, its theoretical underpinnings and limitations, and the methods used in the evaluation of healthcare interventions and programs.
Chapter 3: The economic evaluation of healthcare interventions

Introduction

In the previous chapter the reader was provided with an introduction to Duchenne muscular dystrophy, the natural history of the disease, its management and the alternative treatment pathways for the initiation of long term mechanical ventilation (LTMV) that were examined in the economic evaluation. In this chapter the reader is provided with an introduction to the theoretical underpinnings of economic evaluation and the characteristics of healthcare as a commodity, which resulted in pervasive market failure and the need for the explicit evaluation of costs and benefits of alternative healthcare interventions. The aim of the chapter is to provide the reader with a basis for understanding the methods that were used in the economic evaluation of the use of LTMV in DMD.

In the first section of the chapter, there is an introduction to welfare economics, an analytical framework for assessing the desirability of the alternative allocation of resources for the production and distribution of goods and services, with the aim of maximising social welfare through the efficient allocation of resources. This is followed by an overview of some of the key assumptions on which the theory is built, and how expected utility theory, the standard theory of individual choice in economics, is used within this framework to explain the choices individuals make about their consumption of goods and services within a perfectly competitive market. The different forms of economic efficiency are then described, which is followed by a discussion of the characteristics of healthcare as a commodity that lead to pervasive market failure, the inefficient allocation of scarce healthcare resources, and the need for explicit evaluation. An overview is then provided of the three main forms of economic evaluation, the questions that they can be used to answer, the strengths and weaknesses of each of the methods and their shared limitations. This is followed by an introduction to the methods that are commonly used in the economic evaluation of healthcare interventions. The chapter ends with details of how the findings of economic evaluation are presented, the methods used in uncertainty and sensitivity analysis, and the quantification and presentation of uncertainty through the use of Cost-effectiveness acceptability curves and the net benefit framework.
Welfare economics

Welfare economics is an analytical framework for assessing the desirability of the alternative allocation of resources for the production and distribution of goods and services, with the aim of maximising social welfare through the efficient allocation of resources. The starting point for analysis within this framework is individual utility, with the desirability of the alternate allocation of resources determined from the sum of individual utility. There are a number of key assumptions upon which this analysis is based: (i) individuals are the best judge of their own welfare and act in a rational and logically consistent manner to maximise utility; (ii) the overall welfare of society can be determined from the sum of individual utility; and (iii) individuals maximise their utility through, among other things, their consumption of goods and services in a free and perfectly competitive market (Garber et al, 1996; Hurley, 1998; Olson, Smith & Harris, 1999; McGuire, 2001; Sassi, Archard & Le Grand, 2001; Messonnier & Meltzer, 2003; Drummond et al, 2005).

The choices individuals make about the goods and services they consume within this perfectly competitive market are explained in this framework, in terms of expected utility theory, the standard theory of individual choice in economics. The development of expected utility theory can be traced back to the 1700s and work by Bernoulli about how individuals make decisions under conditions of uncertainty. In Bernoulli’s theory the amount of money that an individual is prepared to pay for a gamble is not equal to the amount they hope to win, but on the subjective value or utility that they place on the expectation of that win. The application of Bernoulli’s work to modern economic theory dates back a book published in the 1940s, by von Neumann and Morgenstern, entitled the ‘Theory of Games and Economic Behaviour’ (von Neumann & Morgenstern, 1947). In this book von Neumann and Morgenstern provided a mathematical proof that could be used to evaluate rational decision-making under conditions of uncertainty and risk. In this proof: (i) a numerical value is assigned to every possible outcome of a decision, one which rates the desirability, or utility, of that outcome relative to all of the other possible outcomes; (ii) a probability is assigned to the outcomes that reflects the likelihood that they will occur if the decision is taken; (iii) the utility of each of the outcomes is weighted by their probability to determine the expected value of that outcome and; (iv) the outcome that returns the greatest expected utility is the one that is followed (Schoemaker, 1982; Garber et al, 1996; Snider, Holtgrave & Dunet, 1996; Starmer, 2000; Meltzer, 2001; Petitti, 2000; Goldie & Corso, 2003; Drummond et al, 2005).
The overall welfare of society, within the welfare economic framework, is determined from the sum of the individual utility. In early neo-classical welfare economic theory, utility was assumed to be cardinal\(^{18}\), expressing both the strength and direction of individual preferences which made utilities interpersonally comparable. In the early 1930s there was a shift towards utility measured on an ordinal scale\(^{19}\), which meant that utilities were longer interpersonally comparable, as utility measured on this scale does not reflect either the strength or direction of an individual's preferences. A further consequence of the move to ordinal utility theory was a shift away from the simple calculation of social welfare as the sum of individual utility, towards a set of criteria known as Pareto Optimality. Vïfredo Pareto was a nineteenth-century sociologist who developed a set of general principles for the evaluation of economic policies that could be used to judge whether they would result in an improvement in social welfare. There were two fundamental concepts within this set of principles, these being: (i) an actual Pareto improvement and (ii) Pareto efficiency. An actual Pareto improvement is when the resources available to society are re-allocated in such a way to make one or more individuals better off, and no other individual worse off, with a resultant overall improvement in social welfare. The relative size of the improvement is not important, or who receives it, but simply that no individual is worse off. Pareto efficiency is achieved when there is no other possible allocation of resources that could produce this outcome, and there is the simultaneous achievement of efficient production and consumption. This occurs when the marginal rate of substitution of the number of units of any two goods by an individual is the same, and there is no difference in the technical rate of substitution, the ability to produce a constant number of units of either of these two goods (Schoemaker, 1982; Garber et al, 1996; Hurley, 1998; Last, 2001; McGuire, 2001; Sassi, Archard & Le Grand, 2001; Tsuchiya & Williams, 2001; Drummond et al, 2005).

There are three general forms of efficiency. The first two relate to the supply and production of goods and services, and the third to consumer demand. Technical efficiency is achieved when the production of a given output is organised in such a way to minimise the use of labour and capital. For any given output, technical efficiency may be achieved through many different combinations of labour and capital. Cost-effective efficiency is achieved when the production of a technically efficient output is obtained at the least cost. Allocative efficiency is achieved when the production and distribution of a given output is organised in such a way that the number of units produced is equal to consumer demand. Allocative efficiency

---

\(^{18}\) A cardinal outcome is one that can be measured using either an interval or ratio scale. An interval scale is where a change of one unit in one section of the scale, is equal to one unit of change in any other section of the scale. Temperature is an example of an interval scale. A ratio scale meets all the requirements of an interval scale, while also containing a true zero point, i.e. weight and height (Garber et al, 1996; Last, 2001; Drummond et al, 2005).

\(^{19}\) An ordinal scale is one that contains distinct ordered qualitative categories for which there is no meaningful numeric difference. For example, an individual may be categorised as being working class, middle class, or upper class, but that is not to say that an upper class individual has three times the value, or worth of a working class individual (Garber et al, 1996; Last, 2001; Drummond et al, 2005).
cannot be achieved unless the production of a given output is organised in a technically efficient and cost-effective manner (Richardson et al, 1998; Carter & Harris, 1999; Hurley, 2000).

In welfare economic theory individuals maximise their utility, at least partly, from their consumption of goods and services within a perfectly competitive market. In this market: (i) all of the goods and services that an individual may wish to buy, are priced and available for purchase; (ii) individual consumers have perfect information about these goods and services, and maximise their utility through the rational and informed choices they make about their consumption; (iii) consumer demand results in the production of a mix of goods and services that equals this demand and; (iv) there is a Pareto optimal allocation of resources due to the presence of competitive equilibrium. A state of competitive equilibrium exists within this market as a result of consumers acting to maximise their utility, and producers acting to maximise their profit. The mix of goods and services, and the quantities produced within this market exactly equals consumer demand, with the prices paid by consumers and the payments received by producers being solely the result of market forces (the independence of supply and demand). There is supply side competition, and these goods and services are produced in a technically efficient manner, with the prices charged by producers accurately reflecting the opportunity cost, the mix of labour and capital used in their production, and the return that could be have been gained from the alternative use of these resources. If there are no increasing returns from production, and every good or service that an individual may wish to consume is priced and available in the market, then this will result in an allocation of resources that is Pareto optimal for any given distribution of purchasing power, as there is no other allocation of resources that would improve the utility of one individual, without decreasing the utility of another. It should be noted, that any market that deviates in any significant way from this perfectly competitive model will by definition not be able to achieve optimality (Arrow & Debreu, 1954; Arrow, 1963; Garber et al, 1996; Jacobs, 1997; Olson, Smith & Harris, 1999; Hurley, 2000; McGuire, 2001).
Healthcare as a commodity

In a perfectly competitive market there is complete certainty, no random events, and information is freely available to all. Consumers know as much about the goods and services in the market, as do the companies that produce them, and they know with relative certainty the benefits that they will obtain from their consumption of these goods and services. This is far removed from the circumstances that surround healthcare, where there is significant uncertainty surrounding the decisions that a consumer needs to make about the purchase of healthcare, for example, in relation to their risk of disease, their need for treatment, and the effectiveness of that treatment. An individual who seeks out the services of a medical, or other health professional, does this to some extent simply to attempt to reduce this uncertainty through the purchase of knowledge. In doing so the consumer, who may otherwise feel well, may find that based on information that is only available to the health professional, such as the results of diagnostic tests, that they are prescribed treatments for problems that they were unaware they had, which while being appropriate, may be expensive, invasive or dangerous (Arrow, 1963; Garber et al, 1996; Jacobs, 1997; Dranove & Satterwaite, 2000; Hurley, 2000; Segal & Chen, 2001).

This difference in medical knowledge, and access to the results of diagnostic and other tests, means that an individual must trust their medical professional’s judgement and allow them to act as their agents, and to determine both the form and the level of their consumption of medical services and treatment. This trust is based on an expectation that the health professional will behave differently than other producers in the market. They are expected to act only in the individuals’ interest, with any plans they make for the consumptions of their services being, clearly and objectively, based only on the results of diagnostic testing and examination, and not driven by self-interest. As the need for trust puts the health professional in a position where they can easily exploit the consumer’s lack of knowledge, and induce them to purchase more of their services, or less of the services provided by other producers, than they may have otherwise chosen to. A medical professional who chooses to do this, violates an explicit assumption of the model of a perfectly competitive market model, that of the independence of supply and demand (Arrow, 1963; Jacobs, 1997; Richardson & Peacock, 1999; Dranove & Satterwaite, 2000; Hurley, 2000).

The supply of medical professionals is limited by licensing and the relatively small numbers of undergraduate and post-graduate training positions that are available in medical schools. This is ostensibly based on a need to maintain quality but it acts to increase the cost of medical care by restricting its supply. While the cost of a medical education is relatively high, only a small proportion of this cost is met directly by students, with the majority of the cost being met by government or private sources. Therefore, the benefits gained from a medical education for a student greatly exceed its cost, which in
theory should act to reduce the cost of production of medical care. In practice the restriction on supply means this has no significant impact on the cost of medical care, and ultimately this is determined by non-market mechanisms. Furthermore in a perfectly competitive market, if a commodity can have different qualities, then these are available for purchase at different prices to suit different tastes and incomes, but again the licensing of medical practitioners and the restrictions placed on training positions limits the forms of medical care that are available, and at what cost (Arrow, 1963).

The onset of illness is an unpredictable random event that has, at the very least, the potential to significantly impact on an individual's quality of life and their ability to earn an income. While there is significant uncertainty surrounding the effectiveness of healthcare, for most individuals the expectation of benefit greatly exceeds the cost of treatment, even when this cost is high in relation to an individual's income. This expectation of benefit, and high cost of healthcare, creates a demand for insurance that is unlike that for any other good or service. While in theory it should be possible to create a market for the risk of random events, and for individuals to obtain insurance that completely covers them against this risk, the uncertainty surrounding the need for healthcare and the effectiveness of treatment for a given individual is so great, that insurance at the required level of complexity and uncertainty is generally not available. The healthcare insurance products that are available in the market generally only meet the cost of medical care, with or without a co-payment from the individual, with restrictions placed on the treatments and periods of hospitalisation etc.; that are covered (Dranove & Satterwaite, 2000).

In a perfectly competitive market there are no positive or negative externalities of either production or consumption. A common example of a negative externality of production is pollution. A firm that generates air or water pollution in the process of maximising their profit creates a negative externality of production, through the imposition of a cost on society in the form of reduced air or water quality. If they are allowed to do this without paying for the cost of the pollution they generate, this results in a less than optimal outcome for society. An example of a positive externality of production is a plantation forest. While the purpose of the forest may be to produce wood for building or furniture, as it grows the forest absorbs carbon dioxide and through doing so produces a benefit for society in the form of an improvement in air quality (Arrow, 1963; Rice & Unruh, 2003).

A negative externality of consumption is where an individual's purchase of a good or service reduces the utility of another individual, with a positive externality of consumption being where the purchases of goods and services by one individual increases the utility of another individual. The smoking of cigarettes in public places is an example of a negative externality of consumption. While the smoker gains the full benefits of their consumption, the utility of non-smoking individuals is decreased as a result of the
degradation of the ambient environment (the air quality in public places) and the health impact of the inhalation of “second-hand smoke” (environmental tobacco smoke). One of the most widely cited examples of a positive externality of consumption in healthcare is that of immunisation. An individual who purchases and meets the full cost of immunisation for a communicable disease reduces not only their own risk of developing that disease, but also the risk of other individuals who have not purchased immunisation. Unfortunately this can result in less than a socially desirable purchase of immunisation, as an individual who chooses not to purchase immunisation for a given disease because of a perceived reduction in risk due to the purchases of others, may go onto develop the disease and in doing so infect others. Ideally an individual who chooses not to purchase immunisation, because of this perceived reduction in risk, should be required to pay for this benefit or for the risk they impose on others. The price they would need to pay for the risk they impose on others would need to be sufficiently high to compensate these individuals for their potential loss of health, or alternatively a payment could be made to the individual to induce them to undergo immunisation. In practice, market forces are unable by themselves to achieve socially desirable levels of public health activities such as immunisation. To achieve socially desirable levels of these activities requires government intervention in the market, such as through the use of incentives such as public subsidies for immunisation, or through the use of compulsion, such as the removal of government benefits or increases in taxation (Arrow, 1963; Rice & Unruh, 2003; Lambert, 2006).

The characteristics of healthcare as a commodity and the presence of significant externalities are such that there is no other substantial market that violates the underlying assumptions of a perfectly competitive market to the extent that healthcare does. It is generally accepted that this results in pervasive market failure and the sub-optimal and inefficient allocation and use of resources. As a consequence of this failure, non-market mechanisms have been developed for use in the healthcare sector, such as economic evaluation, that can be used to explicitly evaluate the costs and benefits of alternative healthcare interventions, policies and programs so that healthcare outcomes can be maximised through the efficient use of resources (Arrow, 1963; Dranove & Satterwaite, 2000; Hurley, 2000; Segal & Chen, 2001; Rice & Unruh, 2003; Savedoff, 2004).

Economic evaluation

There are three main forms of economic evaluation: Cost-benefit analysis (CBA); Cost-effectiveness analysis (CEA) and; Cost-utility analysis (CUA). A common feature in all of these forms of evaluation is the consideration of the inputs (costs) and outputs (benefits) of an intervention, and a comparison of alternative courses of action. The choice of the form of evaluation to use in given analysis is determined
by: (i) the question that needs to be answered; (ii) the information needs of the decision maker; (iii) the outcome measures of interest and; (iv) the availability of data. As an aid to decision-making, economic evaluation can only be used where the objective of an intervention is clearly articulated, and there is at least one alternative course of action against which an assessment of efficiency can be made. If the focus of the evaluation is relatively narrow, for example, a comparison of the relative costs and benefits of two alternative healthcare interventions that produce the same outcome, then CEA may be all that is needed. If the focus of the evaluation is much broader, for example, a decision about which of a set of alternative healthcare interventions that produce a range of different outcomes should be implemented, then CUA may be all that is needed. If the focus of the evaluation is very broad, and it concerns whether or not to use the resources available to society to implement a given healthcare intervention as opposed to the use of these resources in other sectors of the economy (i.e. education and transport), then CBA is the only form of evaluation that can be used (Torrance, Siegel & Luce, 1996; Richardson et al 1998; Carter & Harris, 1999; Hurley, 2000; Farnham & Haddix, 2003; Drummond et al, 2005).

Cost-benefit analysis (CBA)

Cost-benefit analysis (CBA) is derived from welfare economic theory and it has as one of its basic principles, that the valuation of the benefits of an intervention should be determined by individual consumers. It is the only form of economic evaluation in which both the costs and benefits of an intervention are measured in monetary terms, and in circumstances where competitive markets do not exist, or there is market failure, it can be used to mimic the allocation of resources in a perfectly competitive market. The results of a CBA are usually reported in terms of the net-benefit produced by program or intervention, above that of a comparator, which is determined from the incremental benefit (in monetary terms) less the incremental cost, with a decision about whether an intervention should be implemented based on its ability to produce a positive net benefit. The valuation of health outcomes in monetary terms in CBA means that it can theoretically be used to answer questions with a much broader scope than can other forms of economic evaluation. As it can (at least in theory) be used to inform decisions about the allocation of societal resources for healthcare, as compared to the alternative uses of these resources in other sectors of the economy such as transport or education (Garber et al, 1996; Torrance, Seigel & Luce, 1996; Hurley, 1998; Olson, Smith & Harris, 1999; Sassi, Archard & Le Grand, 2001; Farnham & Haddix, 2003; Hauck, Smith & Goddard, 2004; Drummond et al, 2005; McGuire, Henderson & Mooney, 2005; Coast, Smith & Lorgelly, 2008.)
Unfortunately for a given set of public resources there can be more than one allocation that is Pareto optimal, and in such situations it is difficult to determine which of the alternative allocations should be implemented. There are also very few situations where the allocation of public resources does not make at least one individual worse off. To address these problems, there was a move to the use of the less restrictive criteria of a Potential Pareto Improvement (Kaldor-Hicks criteria), where a reallocation of resources is deemed to be efficient if the benefits gained by an individual, or group of individuals, is large enough to enable them to hypothetically compensate for any loss incurred by any other individual, or group of individuals. This compensation for their loss would hypothetically make the individuals, who are the losers of the re-allocation, no worse off than they previously were and as a consequence of this, the reallocation of resources would be able to produce a net increase in social welfare (Garber et al, 1996; Torrance, Seigel & Luce, 1996; Hurley, 1998; Olson, Smith & Harris, 1999; Sassi, Archard & Le Grand, 2001; Farnham & Haddix, 2003; Hauck, Smith & Goddard, 2004; Drummond et al, 2005; McGuire, Henderson & Mooney, 2005; Coast, Smith & Lorgelly, 2008).

There are three approaches that are generally used in CBA for the monetary valuation of outcomes: (i) human capital; (ii) revealed preferences and; (iii) contingent valuation. The human capital approach dates back to work undertaken by Dorothy Rice in 1966 to develop a method that could be used to value the cost to society of morbidity and premature mortality. The human capital approach values a healthcare intervention in terms of the restoration or enhancement of the productive capacity of an individual, with labour force participation rates and the current market rate for wages used to determine the present value of future earnings. This approach is not without its problems, for while in theory the productivity of a worker is reflected in the wages paid to them, there are labour market imperfections that see individuals of different genders paid different amounts for the same job. It is an indirect measure of valuation, and as such it violates one of the tenets of welfare economics, that the value of an outcome is best judged by consumers. Furthermore, it values the life of an individual's solely on the basis of their productive capacity as a function of their health status (Garber et al, 1996; Olson, Smith & Harris, 1999; Haddix, Corso & Gorsky, 2003b; Messonnier & Meltzer, 2003; Drummond et al, 2005).

The revealed preference approach values a healthcare intervention using proxy values for life and health that are present in the market place. There is a well-documented trade-off between risk and wages, where individuals who undertake jobs that have the potential to put their health at risk require that a wage premium be paid as compensation. This premium acts a proxy for an individual's preference for the value they place on their health and the trade-off they are willing to make for the goods and services they can consume with a higher income. The strength of this approach is that it is not based on the preference statements of individuals about hypothetical scenarios but on actual choices made in the market place.
This approach is also not without its problems, as the wage paid to a worker may not reflect the risk associated with a job but a lack of alternative jobs that the worker has the training or ability to undertake. The results of the trade-off between risk and wages may not accurately, or solely represent an informed decision made by a rational individual due to imperfections in the labour market, and the perception of occupational risk, and the reported estimates of the wage premiums paid to compensate for risk vary widely, and are job and context specific. Furthermore, to be able to use this approach in an economic evaluation the health outcome needs to be one about which individuals make occupational choices, and for which there is an existing wage premium in the market (i.e. the wage premium paid to individuals who work with heavy metals, in an evaluation of intervention to reduce workplace exposure to these metals) (Olson, Smith & Harris, 1999; Messonnier & Meltzer, 2003; Drummond et al, 2005).

The contingent valuation approach uses survey methods to present hypothetical scenarios to respondents that require them to think about the maximum amount they would be willing to pay for the benefit produced by a healthcare intervention. For the estimate produced by this approach to be accurate respondents need to assume, that for a hypothetical preventative treatment for a disease that: (i) they would need to meet the full treatment costs of the disease if they caught it, and the wage costs for any time lost from work and; (ii) the amount they estimate they would be willing to pay for the intervention would need to be in addition to any costs they were currently incurring for the treatment of the disease, or to reduce their risk of developing it. As with the other two approaches for the monetary valuation of benefits in CBA the contingent valuation approach is not without its problems as, the amount individual's claim they would be willing to pay for a hypothetical treatment may not accurately reflect the price that they would actually be willing to pay if it was available in the market. There are also many individuals who have private health insurance, or access to government subsidised medical treatment, and consequently the amount they report they would be willing to pay for the intervention may not be a true reflection of its value, but of the fraction of the cost of treatment that they would be required to meet. Furthermore, a respondent who believes they are at risk of the disease, who also believes that the intervention might be effective, may simply agree that they would be willing to pay a specified amount for the intervention without really thinking about their reasons, if the price was less than a threshold (i.e. $ 100) about which they were indifferent (Olson, Smith & Harris, 1999; Farnham & Haddix, 2003; Messonnier & Meltzer, 2003; Drummond et al, 2005).

The valuation of health outcomes in monetary terms in CBA, has been, and continues to be controversial as there are many health economists and decision-makers who are uncomfortable with the need to assign explicit monetary values to the additional years of life produced by an intervention, or to the pain and suffering that an individual may experience. There are others who simply reject, on philosophical grounds,
the ethical proposition that an individual’s willingness to pay for a healthcare intervention is an accurate measure of its value. Notwithstanding, individuals clearly gain utility from additional years of life, with most preferring a longer life to a shorter life, and a lower risk of death to a higher risk. The impact that an intervention has on life-expectancy is also a core measure of the health of the general public, and of the effectiveness of many clinical interventions, and it is therefore clearly an important outcome that needs to be taken into account in an economic evaluation. The difficulties in placing a monetary value on morbidity and mortality in CBA ultimately limited its use in the economic evaluation of healthcare interventions, and led to the development in the 1970s of Cost-effectiveness analysis (CEA). Cost-effectiveness analysis is a less complex form of economic evaluation in which the results are expressed in the natural units of outcome produced by an intervention (Garber et al, 1996; Scheck-McAlearney, Schweikart & Pathak, 1999; Hurley, 2000; McGuire, 2001; Tsuchiya & Williams, 2001; Gift, Haddix & Corso, 2003; Messonnier & Meltzer, 2003; Drummond et al, 2005).

**Cost-effectiveness analysis (CEA)**

In Cost-effectiveness analysis (CEA), while costs are measured in monetary units as in CBA, benefits are measured in the natural units of outcome that are the most appropriate for the intervention under analysis. Natural units of outcome such as: life years saved by chemotherapy; cases of disease prevented by a diabetes prevention program; cases of disease averted by a vaccine or; the number of head injuries prevented by the use of bicycle helmets. The results of a CEA are reported as an Incremental Cost-effectiveness Ratio (ICER), the ratio of the difference in incremental costs divided by the difference in incremental effectiveness, where the cost of intervention minus the cost of the alternative (the net cost) is divided by the incremental effectiveness, the effectiveness of the intervention minus that of the comparator (the net benefit). Cost-effectiveness analysis is used to answer questions of technical efficiency. It is similar to the process that is used by firms to determine which of a set of alternative methods of production, to generate a given level of outcome for a specific good or service, has the lowest cost. The use of natural units of measurement in CEA limits the questions it can be used to answer, as it can only be used in situations where there is a clearly specified single objective against which an assessment of effectiveness can be made. It cannot be used to make direct comparisons about the relative value of interventions that use different units of benefit, for example, whether the benefits gained from the detection of a tumour are greater than those for the prevention of a fracture (Garber et al, 1996; Gold et al, 1996; Mandelblatt et al, 1996; Olson, Smith & Harris, 1999; Hurley, 2000; Farnham & Haddix, 2003; Drummond et al, 2005).
Cost-effectiveness analysis, unlike CBA, has no implicit decision rule about when an intervention should be implemented, as it only returns a net cost per unit of health outcome (i.e. the cost per life saved). It is left to decision makers to judge whether the unit of health outcome has sufficient intrinsic value to justify the cost of achieving it, in relation to the resources available to them and the other uses to which these resources could be put. This can be a relatively straight-forward process when there is a clearly divisible set of interventions, a fixed budget, and the maximisation of health is the only objective. In this situation, all a decision maker has to do is to list the interventions in order by their Cost-effectiveness ratios, from the most cost-effective to the least cost-effective, and implement them in order until the budget is exhausted. In situations where the budget is not fixed, and the implementation of an intervention may require the re-allocation of resources from other sectors of the economy, the results of a CEA cannot in itself be used to determine whether this is worth doing. All a decision-maker can do in this situation is set a threshold value for the ratio of costs to benefits that must be achieved by an intervention before it can be considered for implementation (Gold et al, 1996; Seigel, Weinstein & Torrance, 1996; Olson, Smith & Harris, 1999; Gift, Haddix & Corso, 2003; Folland, Goodman & Stano, 2003; Drummond et al, 2005).

In recent years there has been an increasing emphasis on the need for economic evaluations to take into account not only the impact that a healthcare intervention has on morbidity and mortality, but also on quality of life. This change in emphasis is the result of: (i) a steady increase in life expectancy in modern industrialised nations over the last 100 years; (ii) a change in the disease patterns in these nations, from predominately infectious in nature, to predominately chronic in nature; (iii) an ever increasing number of interventions that can be used to save the lives of the acutely ill, or to prolong the lives of the chronically ill; (iv) an understanding that the life years gained from an intervention may not necessarily be healthy ones and that there may be a degree of ongoing pain or disability that some individuals may not be willing to accept and; (iv) a growing acceptance that health care budgets are limited and that it will not be possible to fund every available intervention and consequently that choices will need to be made (Gold et al, 1996; Hawthorne, Richardson & Osborne, 1999; Dolan, 2001; McGuire, 2001; Tsuchiya & Williams, 2001).
Cost-utility analysis (CUA)

Cost-utility analysis (CUA) is a form of economic evaluation in which the impact of an intervention is measured in terms of the cost per QALY (quality adjusted life year), a summary measure that combines the life-years gained from an intervention with their utility weighted quality of life. The quality of life of an individual is represented in this measure on a scale from 0 to 1, with 0 corresponding to a health state equivalent to death and 1 to a state of perfect health. To calculate the number of QALYs produced by an intervention, the time spent by each individual in a given health state is multiplied by the quality of life that is associated with that health state, and the results are summed. For example, if a year of ambulatory dialysis (for renal failure) is associated with a mean utility weighted quality of life of 0.72, and assuming that there is no change over time, then five years of the intervention would produce a total of 3.6 QALYs. The use of a standardised, directly comparable unit of outcome in CUA means that comparisons can be made between the benefits gained from interventions that are used to treat vastly different conditions and diseases. The intervention that returns the lowest cost per QALY, all else being equal, is the one that offers the best value for money (Garber et al, 1996; Hurley, 2000; Hawthorne, Richardson & Day, 2001; Meltzer, 2001; Folland, Goodman & Stano, 2003; Drummond et al, 2005; McKie & Richardson, 2005).

The QALY framework can be traced back to the work of Fanshel & Bush (1970) and Weinstein and Statson (1976). It has expected utility theory as its theoretical foundation, and based on this theory a QALY is an valid representation of health-related utility given the following assumptions: (i) the health status of an individual, and how this changes over time, can be represented and valued in terms of the length of time spent in various states of health; (ii) interventions that produce the same increase in life-expectancy are equally preferred, regardless of their survival curves; (iii) individuals are willing to trade-off years of life spent in less than perfect health at a constant rate for fewer years of life of perfect health, regardless of the number of years that they would be spend in less than perfect health (the constant proportional trade-off assumption) and; (iv) the quality weights used in the construction of a QALY are based on preferences that are measured on an interval scale that ranges from zero (death) to perfect health (1) with the two extremes of the scale being of equal value to all individuals (Garber et al, 1996; McGuire, 2001; Meltzer, 2001; Drummond et al, 2005; Weinstein, Torrance & McGuire, 2009).

The development of an instrument to measure utility weighted quality of life within the QALY framework is a two-step process. The first step is the construction, testing and validation of a system that describes the health state, or states, of interest. The second step is the determination of the utility values attached to these health states. There are two broad approaches that are generally used to describe the health state, or states, of interest. The first is known as the holistic or composite approach, in which a scenario is
developed that describes in detail the health state or states. These scenarios are developed from data collected from patients through the use of focus groups, or some other form of qualitative data collection. The second approach uses a multi-attribute survey instrument with a series of questions that relate to the health state, or states, of interest with a set of possible responses for each question. This method can be used to describe a broad range of possible health states to which scale values and utilities can be attached. There are three techniques that are generally used for the direct measurement of the preferences of individuals for the health states represented by these descriptive systems: (i) rating scales; (ii) the standard gamble; and (iii) the time trade-off (Richardson et al, 1998; Dolan, 2001; Drummond et al, 2005; Hawthorne, Richardson & Day, 2009).

The representation of the desirability of a health state in the rating scale approach can take the form of a list of categories (i.e. A-F), or numbers (i.e. 0-100). When using a rating scale individuals are told to assume that the scale is linear and that each point on the scale is an equal distance apart. They are asked to use the scale to rate set of health states from their least preferred to their most preferred, on a scale from 0 to 1, or 0 to 100. As an alternative, respondents can be asked to create a scale on which they score the health states of interest, from their least preferred to their most preferred with the distance between the scores corresponding to the difference in their preferences. While doing this they are asked to focus not on the scores themselves but on the intervals between each health state, and how they compare to each other. Regardless of the technique that is used, the aim of the rating scale approach is to generate an interval scale of preferences. The ability of the rating-scale approach to do this is questionable due to: (i) the effects of measurement bias, specifically end-of-scale and context bias, where individuals tend to avoid the ends of scales and space out outcomes across the scale regardless of their preferences and; (ii) there being no requirement in the technique that individuals choose between the health states on the scale, or make any trade-offs to obtain them, as is required by expected utility theory (Dolan, 2001; Dasbach & Teutsch, 2003; Drummond et al, 2005).

The standard gamble technique is based on expected utility theory and is consequently considered by many economists to be the gold standard for the direct measurement of preferences. When it is used, for example, to measure preferences for chronic health states a respondent is offered a choice of two alternatives using a lottery like technique to simulate decisions made under conditions of uncertainty. The first alternative is a treatment that has two possible outcomes about which there is uncertainty that could: (i) return the individual to perfect health for a period of t years (probability p) or; (ii) result in the individual’s immediate death (probability 1-p). The second alternative is a treatment that produces less than perfect health (a chronic disease state) for a period of t years. The probability p is varied up to the point that the respondent is indifferent to which of the two alternative treatments they would receive, with the preference
score for the health state for time \( t \) equal to the final value of \( p \) (Dasbach & Teutsch, 2003; Drummond et al, 2005).

The time-trade off technique is a relatively simple method that is used to determine the maximum number of years of perfect health that an individual is willing to trade-off to avoid having to live in a less desirable health state for the rest of their lives. When used to measure a chronic health state that is perceived as more desirable than death, respondents are offered two alternatives: (i) health state \( i \) for time \( t \) followed by death, with time \( t \) corresponding to the life expectancy of an individual with the chronic disease of interest or; (ii) perfect health for time \( x < t \), followed by death. The time period represented by \( x \) is then varied until the point that the respondent is indifferent to the alternatives, with the preference score for health state \( i \) derived from \( i=x/t \). When it was originally developed the time trade-off technique was believed to produce comparable scores to the standard gamble but based on the findings of theoretical research and empirical studies this was found to be incorrect. The scores produced by the two methods are not directly comparable, and the scores produced by the time trade-off technique must be adjusted before they can be used as utilities (Dasbach & Teutsch, 2003; Drummond et al, 2005).

In circumstances where the maximisation of health within a budget constraint is the only objective of a decision-maker, then the QALY framework provides a simple and easy to understand method that can be used to prioritise the implementation of healthcare interventions. As, all else being equal, the interventions that produce the most QALYs at the lowest cost are those that should be given the highest priority, followed by those that produce the next highest number of QALYs, and so on until the budget is exhausted. As there is no consideration of how these QALYs are distributed, it is in a sense egalitarian, as there is no discrimination or selection of interventions of the basis of the social class, occupation or gender of the individuals receiving the QALYs, or of the disease or condition that is being treated. Unfortunately, in the process of treating all QALYs equally, the simple maximisation of health within the conventional QALY framework can inadvertently discriminate against certain individuals and groups in society, such as the disabled and the elderly, as interventions for these groups may produce less QALYs, at the same cost, than interventions for young, able bodied, healthy individuals (McKie & Richardson, 2005a; Weinstein, Torrance & McGuire, 2009).

There is a limited ability in all of the forms of economic evaluation to take into account social values such as equity and fairness. There is no consideration in economic evaluations, or allowance for factors such as: (i) the severity of the pre-treatment condition or disease; (ii) the potential for the realisation for health; (iii) the concentration or dispersion of health benefits or (iv) the age of the target population. These are issues that of are greater concern to society in general, than the simple maximisation of health-gain within
a resource constraint. There is clear evidence from a number of studies that, all else being equal, interventions for individuals with severe illnesses are valued above those for individuals with less severe illnesses, and that priority should be given to these interventions. There is also extensive evidence that it is important to society, that the potential to benefit from an intervention should not be used to discriminate against the chronically ill or permanently disabled, and that they should be allowed to realize their maximum potential for health, even when this is not great. To do otherwise risks placing these individuals in “double jeopardy” where to add to their misfortune of being chronically ill or permanently disabled, their access to healthcare is restricted due to their not being able to benefit from an intervention to the same degree as an able-bodied individual (Dolan, 1998; Sassi, Archard & Le Grand, 2001; Rice & Unruh, 2003; Drummond, 2004; Drummond et al, 2005; McKie & Richardson, 2005a; McKie & Richardson, 2005b; Weinstein, Torrance & McGuire, 2009).

There is also evidence that, all else being equal, interventions that produce a small benefit for a large number of people are of greater value to society, than interventions that produce a large benefit for a small number of people. The age of the individuals who gain the benefits produced by an intervention is also of interest to society, with interventions that benefit the young being generally seen to be of higher value, than interventions that benefit the elderly. There is also evidence that people believe that individuals with high-cost illnesses, or illnesses that have only “cost-ineffective” treatments should not be refused treatment or abandoned by the healthcare system, simply on the basis of cost. This is based on a belief that all members of society should have equal access to treatment, even if this for no other reason than social solidarity and the preservation of hope. There is also a strong desire within society (what is known as the rule of rescue) to: (i) provide care for identifiable individuals who are at risk of immediate but avoidable death (i.e. solo sailors in long distance yacht races) regardless of the cost of doing so; (ii) to give special consideration to identifiable individuals at immediate risk of death, whether or not this death can be averted and; (iii) to provide care to individuals with debilitating, clearly visible and extremely unpleasant diseases (i.e. Epidermolysis bullosa\footnote{A rare genetically determined disorder that is characterised by extreme susceptibility of the skin and mucosa to separate from underlying tissues in response to minor everyday mechanical trauma, and the formation of bullae (large fluid containing blisters). (Herod, Denyer, Goldman & Howard, 2002; Horn & Tidman, 2002; Schober-Flores, 2003).}) even when the available resources could be better used in the prevention of the deaths of a larger group of individuals in the general population, or for the treatment of a larger number of individuals with less visible diseases (Dolan, 1998; Sassi, Archard & Le Grand, 2001; Rice & Unruh, 2003; Drummond, 2004; Drummond et al, 2005; McKie & Richardson, 2005a; McKie & Richardson, 2005b; Weinstein, Torrance & McGuire, 2009).
The design and analysis of economic evaluations

When designing an economic evaluation a researcher needs to make a series of key decisions about a number of issues, such as form of analysis that will be used (i.e. Cost-benefit, Cost-effectiveness or Cost-utility analysis), the perspective of the study and its time horizon. This is a process known as framing that determines the form the evaluation will take, the usefulness of the results, and how they should be interpreted. The process helps to focus the evaluation on research questions that are relevant to the target audience; it reduces the risk of methodological and analytical errors and helps the evaluation to remain on-track. An important first step in the framing of an evaluation is the identification of the target audience, the decision makers and/or funding bodies that are expected to make use of the results. This is to ensure that the evaluation will be able to produce information that is of value to these groups. To do this an analyst needs to take into consideration: (i) the individuals, group or groups who will participate in the decision making process, and who will make the final decision; (ii) the objective of the analysis, and whether this is to address a specific policy decision as opposed to a general policy discussion; (iii) whether there are any individuals, or groups, who will attempt to influence the decision maker or funding body, for example by undertaking their own analysis, or by providing information directly to the decision maker; and (iv); any guidelines or requirements that a decision maker and/or funding body may have for the conduct and reporting of economic evaluations, that must be adhered to for an evaluation to be accepted (Torrance, Siegel & Luce, 1996; Farnham & Haddix, 2003).

Study question

The second step in the framing of an economic evaluation is the development of a clearly articulated and well-constructed study question. To do this a study question needs to address: (i) the underlying program or policy issues that will be examined in the evaluation; (ii) the alternative intervention strategies, and how they will be delivered; (iii) the comparator that will be used as a baseline; (iv) the analytic method that will be used; (v) the target population; (v) the time frame and perspective of the analysis; and (vi); the outcome or outcomes of interest (Torrance, Siegel & Luce, 1996; Farnham & Haddix, 2003; Drummond et al, 2005).
Selection of the intervention and comparator

The next step in the framing of the evaluation is the selection of an appropriate intervention, or interventions, and a baseline comparator or comparators. The selection of the intervention, or interventions needs to be based on: (i) evidence of a demonstrated link between the intervention and the health outcome of interest; (ii) the available evidence of the effectiveness of the intervention and; (iii) the need to include every appropriate alternative treatment option, while also attempting to keep these to a minimum to prevent the evaluation from becoming untenable. The selection of a valid comparator is an essential component of an economic evaluation, and for any given intervention there are many different comparators that can be used. For example, in an economic evaluation of a lifestyle modification program to prevent the development of Type 2 diabetes mellitus in an at-risk population, the following comparators could be used: (i) an existing intervention (i.e. drug therapy) or competing diabetes prevention program with proven effectiveness; (ii) a no-treatment option or; (iii) the current standard of care provided by general practitioners for at-risk individuals. Once selected, the intervention(s) and comparator(s) need be clearly defined in terms of: (i) the setting in which they will be delivered, and by whom; (ii) the technology that will be used; (iii) the timing of delivery; (iv) the intended target population and; (v) whether or not the intervention will be included as a component of another service (Torrance, Siegel & Luce, 1996; Farnham & Haddix, 2003).

Event pathway

In the next step of the process, the intervention strategies and the comparator are mapped out as detailed event pathways that describe in concrete and well defined steps the clinical events and management activities that may result in the consumption of healthcare resources, and/or produce a change in the health status of the target population. A detailed and complete event pathway is a fundamental component of an economic evaluation that provides a clear link between the events and activities of the interventions and comparator, the health outcomes they produce and the resources they consume. To be able to do this, an event pathway needs to include all of the relevant effects and side effects of the interventions and the comparator, as well as every event that consumes healthcare resources and ideally the timeframe over which these events occur. This requires a level of detail far greater than that of clinical event pathway, for while a move from an acute care hospital to a rehabilitation hospital in a clinical event pathway may not necessarily result in change in health status, it will most likely result in a change in the consumption of healthcare resources (Mandelblatt et al 1996; Torrance, Siegel & Luce, 1996; Donaldson, Mugford & Vale, 2002; Farnham & Haddix, 2003).
Economic evaluations can be conducted from a range of viewpoints, or perspectives, with the choice of perspective being a major determinant of the costs and outcomes that need to be considered in the analysis, and how they are valued. The choice of the perspective is primarily determined by the question that needs to be answered, and the requirements of the decision maker. From the broadest to the narrowest, these perspectives are: (i) societal – all costs and health outcomes, as well as the broader impact of the intervention on other sectors of society (i.e. education, transport), regardless of who gains the benefits or who incurs the costs; (ii) government – all costs that are met by government agencies, such as social services, and the health outcomes achieved; (iii) health sector/third party payer – all costs that are met by health care funding authorities, healthcare insurers, or hospitals/clinics and the healthcare outcomes achieved and; (iv) patient/family – all healthcare and other costs that are met by patients and their families and the health outcomes they gain (Torrance, Seigel & Luce, 1996; Farnham & Haddix, 2003; Fox-Rushby & Cairns, 2005).

Economic evaluations of healthcare interventions should ideally be conducted from a societal perspective as the resources available to society are limited. The use of these resources for healthcare comes at the cost of the loss of the opportunities that could have been gained from their use in other sectors of society, where they could have potentially produced a greater increase in overall social welfare. It is only through the use of a societal perspective that an economic evaluation can fully capture all of the resource use implications of an intervention, including the impact of cost-shifting between sectors of society. For as often occurs in healthcare, costs may be shifted from a hospital to a healthcare insurer or to government, or from an insurer to a patient or other party. For example, while a program for the early discharge of surgical patients from hospital to home could significantly reduce the cost of surgery to the hospital, this could increase the overall cost to society by: delaying the recovery of the patient and consequently the time they spend out of the workforce, increasing the productivity loss to society and; increasing the patient’s use of community clinics, district nursing and informal care (Russell et al, 1996; Torrance, Seigel & Luce, 1996; Byford & Raftery, 1998; Farnham & Haddix, 2003; Johannesson et al, 2009; Jonsson, 2009).

Notwithstanding, the use of another perspective in an economic evaluation such as that of a health-sector/third party payer, may be necessary when a decision needs to be made by about the relative efficiency of alternative health care interventions within a fixed budget. In broad terms the ultimate aim of economic evaluation is to provide decision makers with evidence about the efficient use of resources. This means that an evaluation needs to be conducted from a viewpoint that produces a result that enables
decision-makers to do this, within the constraints placed upon them by their funding bodies (i.e. healthcare insurers and/or government). While the use of a perspective, other than societal perspective in an economic evaluation is less than ideal, the potential non-health benefits of a healthcare intervention (i.e. an associated productivity gain) may be of little value to a decision-maker whose budget is strictly allocated for the maximisation of health, and it is unlikely that evidence of a non-health outcome could result in an increase in their budget from the re-allocation of resources from other sectors of the economy (i.e. education or transportation) (Torrance, Siegel & Luce, et al, 2009; Jonsson, 2009).

**Target population**

The target population of an economic evaluation is the group, or groups, who are directly or indirectly affected by an intervention. This may be a group of individuals who have a given disease, or who are at risk of developing a given disease, or who are a given age and/or gender, or who live in a given region or environment, or any combination of these characteristics. The selection of the target population is an important component of the framing of an evaluation, one that can have a major impact on the Cost-effectiveness of an intervention. For example, a screening program for permanent congenital hearing loss could have as its target population, newborns with bilateral hearing impairment, or newborns with bilateral or unilateral hearing impairment. As the prevalence of bilateral hearing impairment is lower than that of unilateral hearing impairment, a program that screened for both forms of permanent congenital hearing impairment would be able to detect more cases of disease than one that only screened for bilateral hearing impairment, which in itself could act to reduce the cost per case detected. In most circumstances the choice of target population is determined by the study context, for example in an economic evaluation of alternative treatments for hypercholesterolemia, the target population would clearly need to be individuals who had previously been diagnosed with this condition (Torrance, Siegel & Luce, 1996; Davis et al,1997; Gift, Haddix & Corso, 2003; Goldie & Corso, 2003).

**Time horizon**

The time horizon of an economic evaluation is the period of time across which costs and outcomes are tracked. There are two components to a time horizon: (i) the time frame and; (ii) the analytic horizon. The time frame is the period of time during which the alternatives under examination are applied. The time frame of an evaluation needs to be sufficiently long to take into account factors such as: (i) start-up costs; (ii) the cost of ongoing maintenance; (iii) variations in costs and outcomes across holiday periods and seasons and; (iv) the time needed for the intervention to achieve a steady state (a stable level of output). The analytic horizon of an evaluation is the period of time across which all important costs and outcomes
are tracked. It is important that an evaluation is conducted with an appropriate time horizon, one that takes into account all of the possible costs and consequences of the intervention under examination, intended and unintended, as this can have a major impact on the findings of the study. The benefits produced by an intervention may continue for many years, and important life-saving effects may not be observed until far into the future. Consequently, it is not uncommon for the analytic horizon of an economic evaluation to extend for many years into the future, or for the lifespan of the target population, especially when the outcome produced by an intervention is an increase in life expectancy (Gold et al, 1996; Mandelblatt et al, 1996; Torrance, Seigel & Luce, 1996; Brouwer, Rutten & Koopmanschap, 2001; Farnham & Haddix, 2003; Drummond et al, 2005; Fox-Rushby & Cairns, 2005; Griffin et al, 2008; Caro et al, 2012).

The identification of outcomes

A three step process of identification, measurement and valuation is used to determine the outcomes produced by an intervention, or interventions and its comparator. The identification of the outcomes of interest in an economic evaluation is guided by: (i) the policy question that needs to be answered; (ii) the form of economic evaluation that is being used (i.e. CBA, CEA or CUA); (iii) the outcomes that can reasonably be expected to change as a result of the intervention; (iv) the possible side-effects, and other unintended effects of the intervention; (v) the outcome data that it will be possible to collect and; (vi) the outcomes that are of interest to society in general, and to groups within society such as decision makers, clinicians, patients and their families. A common objective of many healthcare interventions is a reduction in morbidity and mortality and it is therefore not surprising that life-expectancy is widely used as an outcome measure in economic evaluations. However for some interventions it can be many years, if not decades, before any change can be observed in the natural history of a disease that results in decrease in mortality, such as a reduction in the number of cases of lung cancer produced by a smoking cessation program. In some circumstances it may be possible to use changes in short term outcome measures, such as smoking cessation rates, as intermediate/ surrogate outcome measures from which long term outcomes can be reasonably predicted (Gold et al, 1996; Seigel, Weinstein & Torrance, 1996; Farnham & Haddix, 2003; Yabroff & Mandelblatt, 2003; Drummond et al, 2005; Centre for Reviews and Dissemination, 2008).

The use of intermediate outcome measures in economic evaluations is to provide decision makers with evidence that they can use to make timely comparisons about the predicted long-term outcomes produced by competing programs that have the same long-term objectives. When an intermediate outcome measure is used in an economic evaluation it must be demonstrated that it has clinical relevance, or is of
value in its own right. For an intermediate outcome measure to be valid and reliable, it needs to have biological plausibility and to be on the casual pathway between the intervention and the final outcome. It needs to capture all of the important effects of the intervention on the final outcome, and be supported by good research evidence for a previously established link to the final outcome. This should be in the form of epidemiological studies that demonstrate that the final outcome can be reliably predicted from the intermediate outcome, and from clinical trials that demonstrate that an intervention that impacts on the intermediate outcome has a corresponding impact on the final outcome. Notwithstanding, economic evaluations should ideally be conducted using patient relevant outcome measures, and not intermediate outcome measures, for while a reduction in blood pressure may be of clinical value, it is of less relevance to a patient than a final outcome such as the reduction in the risk of a stroke, or of a heart attack (Centre for Reviews and Dissemination, 2008; Griffin et al, 2008; Torgerson & Torgerson, 2008; Drummond et al, 2009; Taylor & Elston, 2009).

The measurement and valuation of outcomes

Evidence of effectiveness

The evidence for the effectiveness of an intervention for use in an economic evaluation is usually obtained from the medical literature. There are three issues that need to be assessed in relation to the use of this evidence: (i) quality; (ii) relevance and; (iii) comprehensiveness. The quality of the available evidence is generally assessed using the guidelines that have been developed for use in evidence based medicine. While the best evidence for the size of a treatment effect is obtained from a well conducted randomized controlled trial, in many cases the data required for an economic evaluation must also be obtained from studies that use other research designs (i.e. cohort studies), as this may be the only source of data for the probability of rare or unusual events, such as drug interactions. In practice there are relatively few economic evaluations that are based entirely on the evidence obtained from randomized controlled trials (Drummond et al, 2005).

There is a potential for conflict between the need for high quality evidence of the effectiveness of an intervention for use in an economic evaluations, and the relevance of that evidence. Randomized controlled trials are usually conducted to produce evidence of the efficacy of an intervention by demonstrating that a clear causal relationship exists between an intervention and an outcome. To be able to do this, efficacy studies are typically conducted in academic research centres under ideal conditions, in which intervention is driven by a strict treatment protocol and delivered to a highly selective study population who are closely monitored to ensure compliance with treatment. There is also a higher degree
of interaction between the study participants and medical and other health professionals than would normally occur in everyday clinical practice. Consequently the findings of efficacy studies may not be an accurate indication of the impact that an intervention may have when it is used in everyday clinical practice and so ideally economic evaluations should instead be based on the findings of effectiveness studies. The aim of an effectiveness study is to determine the impact that an intervention has on health outcomes when it is used in everyday clinical practice, where the age, gender and health status of patients can differ significantly from the select group of individuals that are recruited into efficacy studies. Furthermore, even when there are clinical practice guidelines for the delivery and monitoring of an intervention, a healthcare provider may not be aware of the guidelines, or may only partially adhere to them due to time, resource or other constraints. Patients may also decide not to follow medical advice, or be unable to adhere to the requirements of a complex program of treatment (Mandelblatt et al, 1996; Jacobs, Hailey & Jones, 2003; Yabroff & Mandelblatt, 2003; Drummond et al, 2005; Furher, 2007).

The relevance of the findings of a study for a given economic evaluation need to be judged in relation to the decision making context, with consideration given to the setting in which the study was conducted (i.e. an acute care hospital versus a general practice setting) and the expertise of the service providers (i.e. specialist medical practitioners versus general practitioners). While the best available evidence for an intervention should be used in an economic evaluation, in practice this can be difficult to determine, as it is not unusual for studies to produce conflicting results, with some studies finding that an intervention is effective and others finding that it is not. The conflicting results produced by these studies may be due to differences in research methodology, bias, misinterpretation, or misrepresentation. An assessment of the comprehensiveness of the evidence for an intervention, for use in an economic evaluation, needs to take into account the extent to which it is an accurate reflection of the entire research base of an intervention, and not just the findings of a select sub-group of studies. Ideally the effectiveness of an intervention, for use in an economic evaluation, should be obtained from a formal systematic literature review (Mandelblatt et al, 1996; Jacobs, Hailey & Jones, 2003; Yabroff & Mandelblatt, 2003; Drummond et al, 2005; Furher, 2007).

A formal systematic literature review is a process that is followed to identify and evaluate the evidence for an intervention from all relevant studies, with the aim of producing a pooled summary measure of effectiveness. The process is guided by a series of explicit methodological principles and requirements for the conduct and reporting of literature searches, the inclusion and exclusion of studies, the assessment of bias, the use of statistical techniques and other issues. When meta-analysis\(^\text{21}\) is used to quantitatively

\[^{21}\text{The statistical synthesis and pooling of data from a group of separate but similar studies of a intervention to produce a quantitative summary of the overall trend of the results. In most cases, new health-care interventions}\]
combine the results of multiple studies, a summary measure of effect can be produced that is more reliable than a measure of effect that is obtained from a single study. A well conducted formal systematic literature review is consequently able to produce a reliable and defendable estimate of the effect of an intervention. It is also able to demonstrate where there are gaps in what is known about an intervention, and the disease or condition that it is used to treat, and where further research is needed (Dimichelli, Jefferson & Vale, 2002; Drummond et al, 2005; Hanratty et al, 2007; Centre for Reviews and Dissemination, 2008).

produce only modest improvements in patient outcomes. Small scale trials of these interventions may not be able to reliably produce results that reach statistical significance. The pooling of studies in a meta-analysis, increases statistical power and precision, which may result in the detection of a real and statistically significant result. A twostep approach is used in most meta-analyses: (i) summary statistics are calculated for each individual study for the outcome or outcomes of interest and (ii) the summary statistics of the individual studies are converted to a common measures of effect (i.e. Risk ratios (RR) , Odds ratios (OR) & Hazard ratios (HR) which are defined below) and combined to produce an summary estimate of effect, which in most cases is the weighted average of the individual study estimates. With studies commonly weighted in relation to their same size (in inverse proportion to their variance or standard error squared), so that larger studies are given more weight than smaller studies (Last, 2001; Centre for Reviews and Dissemination, 2008).

Risk ratio: An indicator of the change in risk produced by an intervention/event/activity. It is calculated as the probability of an event in an intervention group divided by the probability of an event in a comparator group (the ratio of risk in the two groups). The probability of the event for each group is obtained by dividing the number of events by the number of individuals in the group. A risk ratio of 0.75 is a reduction of risk of three quarters, a risk ratio of 2.0 is a increase of risk of twice that of the comparator group (Last, 2001; Yabroff & Mandelblatt, 2003; Centre for Reviews and Dissemination, 2008).

Odds ratio: The ratio of the probability of an event happening, divided by the probability of it not happening. For example, 25 cases of disease develop in an at risk population of 100 individuals, the odds of this happening are 33% (25/(100-25)). An Odds ratio can be used where the measurement of exposure precedes disease, where this occurs simultaneously, and where the measurement of disease precedes the measurement of exposure (Yabroff & Mandelblatt, 2003). RRs and ORs are different measures of association that cannot be used interchangeably. When the occurrence of an event is rare (i.e. less than 10%) they will return approximately the same results but when the occurrence of a event is high, they will return very different results. An intervention that increases the likelihood of an event has a larger OR than RR, with the reverse for an intervention that reduces the risk of an event. If an OR is interpreted as an RR the effect of an intervention will be overestimated (a quite common error in published reports and systematic reviews). To avoid this error, it may be worth transforming ORs to measures of risk (as a description of change) as these are less likely to be misinterpreted (Centre for Reviews and Dissemination, 2008).

Hazard ratio: For time to outcome events (i.e. survival data from Kaplan Meier analysis). A summary measure of whether or not an event happened/and when it happened. For example, where an intervention for a chronic disease may not be able prevent an event from happening, but only to delay its onset (Centre for Reviews and Dissemination, 2008).
Health-related quality of life

Health-related quality of life (HRQoL) is a broad multi-dimensional concept that encompasses the physical, psychological and social domains of health that impact on an individual's ability to function, and to derive satisfaction from their life. As the quality of life of an individual cannot be directly observed or measured, it is instead measured indirectly through the use of standardised instruments that are valid and reliable, and developed in accordance with the principles of item-measurement theory. The basis of item measurement theory is the proposition that outcomes, such as quality of life, can be measured indirectly through a series of questions (items) that address a set of constructs or concepts for a series of health-related domains. It is also proposed that the responses provided by individuals to these questions can be converted to numerical values from which domain scores can be derived, which in turn can be combined to produce an overall score that represents the quality of the life of an individual (Testa & Simonson, 1996; Abresch, Seyden & Wineinger, 1998; Brazier et al, 1999; Yabroff & Mandelblatt, 2003).

There are three different types of instruments that can be used to measure HRQoL in an economic evaluation: (i) general health profiles; (ii) disease or condition specific questionnaires; and; (iii) preference based multi-attribute utility scales. General health profiles such as the Short Form (SF 36), the Nottingham health Profile, and the Sickness Impact Profile are broad measures that have been designed to encompass a wide range of universally important health-related domains that are not age, gender or disease specific. Consequently, they can be used to make comparisons about the impact of a health care intervention across a wide range of conditions and population groups. However their generic nature means that they may lack sensitivity and responsiveness and consequently they may not be able to detect the small changes in disease specific health-related domains that may be produced by an intervention or treatment. The majority of these instruments have not been developed to produce a single overall index score for the quality of life of an individual, but a simply set of scores for each of the health-related domains. While this produces a profile from which comparisons can be made between the scores for the different domains, the lack of a single overall index of HRQoL limits their use in economic evaluations. These instruments are also not calibrated on a scale from 0 to 1, and are not utility weighted and consequently cannot be used for the calculation of QALYS (Testa & Simonson, 1996; Hawthorne, Richardson, Osborne & McNeil, 1997; Brazier et al, 1999; Dolan, 2000; Janssens, 2001; Yabroff & Mandelblatt, 2003; Drummond et al, 2005; Rajmil, Perestelo-Perez & Herdman, 2010).

22 Demonstrated to truly measure the phenomenon of interest (i.e. anxiety; depression) that it was developed to measure (Mandelblatt et al, 1996; Brazier et al, 1999; Last, 2001).
Disease, or condition, specific questionnaires are designed to measure HRQoL in individuals with a specific disease or medical condition. They do this by focusing on the health-related domains that are the most severely affected by the given disease or condition and the characteristics of the individuals or groups (i.e. age, gender) in whom the disease or condition most commonly occurs. For example, a disease-specific instrument for children with juvenile idiopathic rheumatoid arthritis would most likely include items relating to pain, mobility and age-specific social functioning. These instruments generally have greater sensitivity and responsiveness, than generic measures, to clinically significant changes in HRQoL that may be produced by a healthcare intervention for a specific disease or condition. The responsiveness of these instruments, their ability to detect change, makes them extremely useful in economic evaluations. Furthermore, as the focus of these instruments is on the domains most severely affected by a disease or medical condition they are generally better accepted by study participants and their treating physicians, as they are seen to be more relevant than generic HRQoL instruments. But as they are not comprehensive measures of HRQoL they cannot be used to make valid comparisons of the Cost-effectiveness of healthcare interventions across a range of diseases and conditions (Testa & Simonson, 1996; Brazier et al, 1999; Dolan, 2000; Janssens, 2001; Yabroff & Mandelblatt, 2003; Drummond et al, 2005; Rajmil, Perestelo-Perez & Herdman, 2010).

23 Demonstrate to consistently return the same results (with a minimum amount of random error) when used repeatedly under identical conditions in the same individual, or group of individuals (Mandelblatt et al, 1996; Brazier et al, 1999; Last, 2001).
As previously discussed, the direct measurement of preferences and the calculation of utilities for the multitude of possible health states that an individual may experience is a complex and time consuming process. A widely used alternative is the use of pre-scored multi-attribute utility scales (MAUs) such as: the QWB - Quality of Well-Being (Kaplan & Anderson, 1988); the HUI-III - Health Utilities Index, version 3 (Horsman et al, 2003); the EQ-5D - EuroQol, 5 dimensions (EuroQol group, 1990); the SF-6D - Short Form 6D (Brazier et al, 2002); the 15-D (Sintonen & Pekurinen, 1993) and; the AQoL - Assessment of Quality of Life (Hawthorne et al, 2001). All of these instruments produce a single index score on a scale from 0 to 1 (a health state equivalent to death, through to one equivalent to perfect health) that can be used to calculate quality-adjusted life-years (QALYs) for use in a Cost-utility analysis. While the instruments vary in terms of: (i) how they conceptualize HRQoL; (ii) their descriptive systems; (iii) the health-related domains they measure; (iv) the methods they use to elicit preference scores; (v) the population, or population sub-group in which these preferences were measured and; (vi) the methods that are used to convert item and domain scores into an single index score (as demonstrated in Table 7) but they nonetheless generate remarkably consistent utility scores (Hawthorne et al, 1997; Brazier et al, 1999; Hawthorne, Richardson & Day, 2001; Drummond et al, 2005).

There are a number of factors that need to be considered in the selection of a HRQoL instrument for use in an economic evaluation, such as: (i) whether it has demonstrated reliability, validity and responsiveness in the target population; (ii) the time frame of the response items, i.e. the past month or the past week; (iii) the cognitive burden it places on respondents, as an excessive cognitive burden can lead to fatigue and a reduction in data quality; (iv) if there is a cost associated with its use; (v) the population from which preferences were obtained; (ii) the method that was used to elicit preferences (i.e. a rating scale, or the standard gamble) and; (vii) the instrument is acceptable to respondents, clinicians, ethics committees and decision-makers (Mandelblatt et al, 1996; Brazier et al, 1999; Drummond et al, 2005; Griffin et al, 2008; Rajmil, Perestelo-Perez, Herdman, 2010).

---

24 There are two issues in relation to this question: (i) whether an instrument that uses preferences obtained from a population in one country (i.e. the QWB which uses preference weights obtained from residents of San Diego, USA) can be used in an economic evaluation in another country, for example Australia. The majority of studies that have been conducted to date have shown that when the same method of preference measurement is used, there is little or no difference in the scores obtained in between population groups in different countries and; (ii) the population group or sub-group from which the preferences were obtained. For example, whether they obtained from the general population, or a group of health professionals, or individuals with the disease or health-state of interest. The consensus view is that preferences should be obtained from the general population, for in countries with universal healthcare systems the general population funds the healthcare system through taxation, and consequently they should be able to determine to some degree how this money is spent, with healthcare resources directed towards the treatment of health-states and diseases that they perceive as the least desirable (Dolan, 2001; Drummond et al, 2005).
Table 7: Multi-attribute utility scales

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Items</th>
<th>Domains</th>
<th>Preference Elicitation&lt;sup&gt;25&lt;/sup&gt;</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>QWB</td>
<td>27</td>
<td>Mobility, physical activity &amp; social functioning</td>
<td>RS</td>
<td>American</td>
</tr>
<tr>
<td>HUI-III</td>
<td>15</td>
<td>Vision, hearing, speech, ambulation, dexterity, emotion, cognition &amp; pain</td>
<td>RS transformed to SG</td>
<td>Canadian</td>
</tr>
<tr>
<td>EQ-5D</td>
<td>5</td>
<td>Mobility, self-care, usual activities, pain &amp; discomfort, anxiety &amp; depression</td>
<td>TTO</td>
<td>British</td>
</tr>
<tr>
<td>SF-6D</td>
<td>12</td>
<td>Physical functioning, role limitation, social functioning, pain, mental health &amp; vitality</td>
<td>SG</td>
<td>British</td>
</tr>
<tr>
<td>15-D</td>
<td>15</td>
<td>Mobility, vision, hearing, breathing, sleeping, eating, speech, elimination, usual activities, mental function, discomfort &amp; symptoms, depression, distress, vitality &amp; sexual function</td>
<td>RS</td>
<td>Finnish</td>
</tr>
<tr>
<td>AQoL1</td>
<td>15</td>
<td>Illness, independent living, social relationships, physical senses &amp; psychological well-being</td>
<td>TTO</td>
<td>Australian</td>
</tr>
</tbody>
</table>


<sup>25</sup> RS: rating scale; SG: standard gamble; TTO: time trade off
In the process of designing a study to measure the impact that an intervention has on an outcome such as health-related quality of life, a researcher needs to develop a testable hypothesis, or theory about the effect that the intervention could have on this outcome. For the purposes of this discussion, in statistical theory there are two possible alternative outcomes for any given hypothesis: (i) the null hypothesis, that there will be no difference in the outcome achieved by a group treated with an intervention, compared to that of an equivalent untreated group and; (ii) the alternative hypothesis, that there will be a difference in the outcome achieved in the group treated with an intervention, compared to that of the untreated group. The aim of a study to test a given hypothesis is not to prove the alternative hypothesis but to disprove the null hypothesis. As it will never be possible to demonstrate with 100% certainty that the alternative hypothesis is true but only that there is sufficient evidence of an effect to be able to reject the null hypothesis. To be able to do this, a study must be able to detect a statistically significant change in an outcome between the two groups, or over two points in time, as opposed to the change in this outcome that could occur simply due to chance. This requires that a sample size calculation be performed in the planning stages of a study to determine the number of study participants that will be needed to be able to detect a statistically significant difference (Oakes, 2008; Salanitro, Estrada & Allison, 2008; Torgerson & Torgerson, 2008).

There are two common errors that occur in research studies; (i) when it is inferred that a relationship exists between an intervention and an outcome and the null hypothesis is rejected, when in fact there is no relationship (a Type I error) and; (ii) when it is inferred that a relationship does not exist between an intervention and an outcome and the null hypothesis is accepted, when in fact there is a relationship and the alternative should have been accepted (a Type II error). A Type I error can occur when a $p$ value of 0.01 is interpreted as demonstrating that an intervention produces a strong, statistically significant effect on an outcome, when this is simple due to the sample size of the study. The probability of the occurrence of an event (its $p$-value), is a product of the frequency of the event and the size of the sample, and if an event occurs frequently enough an intervention that has only a small trivial effect on an outcome can appear to produce a statistically significant result. A Type II error can occur when a $p$-value of less than 0.05 is interpreted as indicating that an intervention has no effect on an outcome, when in fact it has a strong effect, but the sample size of the study was not large enough to be able to detect this. The probability of a difference being simply due to chance, for any given effect size, decreases as the sample size increases. The smaller the effect that needs to be detected, the larger the sample size that will be required be able to reliably detect a true effect, if one in fact exists (Oakes, 2008; Salanitro, Estrada & Allison, 2008; Torgerson & Torgerson, 2008).

26 The probability that the result, or one more even extreme could have occurred simply due to chance (i.e. 0.05, 5% or less) (Mullen, 1989; Last, 2001).
A sample size calculation takes into account: (i) errors in measurement; (ii) naturally occurring variations in the outcome measure and; (iii) the risk associated with incorrectly inferring that the intervention has a significant effect on an outcome when in fact it does not (a Type I error), or that it has no effect on an outcome when it fact it does (a Type II error). In studies conducted in human populations sample size calculations are commonly undertaken so that the probability of committing a Type I error is 0.05 (5% or less) and the probability of committing a Type II error is 0.20 (20% or less, a power of 0.80). This ensures that until there is strong evidence for the effect of an intervention that the null hypothesis is accepted, on the basis that it is better to return a false negative result than a false positive result (Oakes, 2008; Salanitro, Estrada & Allison, 2008; Torgerson & Torgerson, 2008).

There are two strategies that are commonly used in the interpretation of the scores produced by HRQoL instruments: (i) the magnitude of the change, where the mean pre & post scores and standard deviations for the two groups are used to calculate an effect size such as Cohen's D\(^{27}\) and; (ii) the minimally important difference, the smallest degree of change (i.e. 0.03) that is believed to be of value to patients, clinicians and decision-makers. These two methods are not without their limitations as: (i) the magnitude of change represented by an effect size is a reflection of the change in the variability of the two samples, as opposed to a change from one health state to another and; (ii) the validity of the calculation of a minimally important difference is reliant on a retrospective self-assessment of a perceived change in the quality of life of a group of individuals (Mullen, 1989; Fletcher et al, 1992; Hawthorne & Osborne, 2005; Rajmil, Perestelo-Perez & Herdman, 2010).

\(^{27}\) The results of this are interpreted as indicating: no change - less than 0.2; a small change - 0.2 to 0.5; a moderate change - 0.51 to 0.8; a large change - 0.8 and higher (Mullen, 1989; Fletcher et al, 1992).
The identification of resource usage

The costing of the resources consumed by a healthcare intervention is a three step process of identification, measurement and valuation. In the first of these three steps, the event pathway that was developed in the early stages of the evaluation is used to map out a detailed and exhaustive list of the healthcare resources consumed by the intervention and the comparator (e.g. labour, capital and supplies/consumables). The event pathway is also used to determine the potential impact that the intervention has on the natural history of the disease, the resultant changes in health status of the target population and the short and long-term resource use implications of these changes. The resultant list of resource usage should, at least initially, include every item of resource use no matter how small, or how difficult it is to measure or to value. When the list is complete, a considered decision is made about the items of resource use that can be reasonably excluded from the analysis, due to their being small or non-essential items that are unlikely to impact on the Cost-effectiveness of the intervention (Luce et al, 1996; Brouwer, Rutten & Koopmanschap, 2001).

The measurement and valuation of resource usage

The measurement of resource usage can be undertaken using a micro-costing approach, a gross-costing approach, or mix of the two broad approaches. In an evaluation that uses a micro-costing approach every single item of resource that is consumed by an intervention is counted, and costed in minute detail, i.e. every blood test, x-ray, drug administered and episode of nursing and medical care provided during a hospital admission. The detailed and precise estimates produced by this approach provides a basis for understanding the relationships between the specific steps of the intervention, their costs and importance, and how these relationships might change as a result of economies of scale. In an evaluation that uses a gross-costing approach an intervention is broken down into as a series of less detailed intermediate products, such as a hospitalisation for a cerebral-vascular accident or a myocardial infarction, that are costed using the average figures for the reimbursement of these events. While the cost estimates produced by a gross-costing approach are less precise than those of a micro-costing approach, they are more generalizable and less influenced by institutional and geographic differences in the treatment of events, and other factors that may impact the use of resources (Luce et al, 1996; Brouwer, Rutten & Koopmanschap, 2001).

A decision about which approach to use in an economic evaluation is ultimately determined by; (i) the available data; (ii) whether or not it will be possible to collect patient level data, as this is a costly, difficult
and time consuming process; (iii) the level of precision required and; (iv) the need for the results to be generalizable. It is not uncommon for a mix of two approaches to be used within a single evaluation, with gross-costing used for items of resource use that occur far into the future\(^{28}\), or for items of resource use that are unlikely to vary much between alternatives, and micro-costing used for items of resource use that occur in the present, or that are likely to vary significantly between alternatives. For example, in an evaluation of a smoking cessation program in which the impact of different combinations of counselling and pharmacotherapy are examined, micro-costing could be used to measure the differences in the costs of counselling and pharmacotherapy, and gross-costing used to estimate the future cost savings that could be obtained from a reduction in chronic obstructive airway disease, cardiovascular disease, lung cancer etc.; (Luce et al, 1996; Brouwer, Rutten & Koopmanschap, 2001).

The final step of the process is the valuation of the resources used by the interventions and the comparator. This is a relatively simple procedure where the units of resource use are multiplied by their cost. This needs to be undertaken on the basis that the consumption of healthcare resources, such as pharmaceuticals, diagnostic testing and medical and nursing time to produce a given healthcare outcome, come at the cost of the loss of the opportunities that could have been produced by the next best alternative use of those resources (the opportunity cost). In most cases the opportunity cost of an intervention is not the same as the financial cost of the outlay of money (i.e. salaries, consumables) that is required to produce it. The opportunity cost of an activity takes into account all of the inputs required to produce a given output, regardless of whether or not they are associated with a direct financial outlay, and includes for example: the time spent by volunteers in a newborn hearing screen program or; the rent-free or subsided office space in a government building that is used by smoking cessation program. While the prices of goods and services in a perfectly competitive market reflects the opportunity cost of their production, the presence of pervasive market failure in the healthcare sector is such, that market prices may not be an accurate reflection of their opportunity cost. The price charged for goods and services, may be higher, or lower, than the opportunity cost, for example when: (i) the price charged by a private hospital for a surgical procedure such as a hysterectomy, or the delivery of a newborn is greater than the costs of the inputs used in their production, even after some allowance is made for a reasonable rate of return on capital or when; (ii) the price charged by a private hospital is less than the cost of production, due to the presence of financial mechanisms such as lump sum financing or the separation of investment costs (Gold et al, 1996; Luce et al, 1996; Johnston et al, 1999; Brouwer, Rutten & Koopmanschap, 2001; Haddix, Corso & Gorsky, 2003a; Drummond et al, 2005; Glick et al, 2007).

\(^{28}\) Less precision is needed in the estimation of future costs, than for present days costs, as there is inherent uncertainty in any estimate of future resource usage. As technological change is inevitable and how this will impact on treatment costs is difficult if not impossible to predict, and even at low discount rates (i.e. 3\%) the impact of small differences in cost decline rapidly over time (Luce et al, 1996).
The resources consumed in the delivery of health care interventions can be categorized as capital or operating costs, and as fixed or variable costs. Capital costs are costs that are incurred at a single point in time for the purchase of assets that have a useable life of greater than one 1 year, such as a magnetic resonance imaging machine. Operating costs are costs that are incurred during a fixed period of time that usually corresponds to the budgetary period of an organisation, for example; the annual cost of the bed linen consumed by a medical ward. These categories can be further broken down into fixed and variable costs based on the level of production of an output. Fixed costs are costs that remain constant in the short term (i.e. over a period of a year) regardless of the level of production of a given output, for example; equipment leases, rent, the design and production of advertising material, and some wages and salaries. Variable costs are costs that vary directly in relation to the level of production, for example, the consumption of bed linen in a medical ward, which increases in relation to the number of patients treated in the ward. Costs such as utilities (i.e. gas & electricity) and administration services (i.e. finance and human resources) are generally treated as fixed costs in an economic evaluation, as they are not directly allocated to a single defined program or service (Luce et al, 1996; Haddix, Corso & Gorsky, 2003a; Drummond et al, 2005).

It is not unusual for the costs of an intervention, and the outcomes it produces to occur at substantially different points in time. So that meaningful comparisons can be made from the same point in time in economic evaluations, costs and outcomes are reported in terms of their net present value. This is undertaken through adjustments that are made for the effects of: (i) inflation, the general increase in the cost of goods and services over time and; (ii) time preference, the general preference of individuals and society for consumption in the present as opposed to the future, and for the incurring of costs in the future, as opposed to the present. The effects of inflation must be adjusted for in an economic evaluation when the cost data that is used for the valuation of resource use is obtained from different points in time, where this difference is longer than one year. For while cost data for a given reference year, i.e. 2012 can be summed without adjustment, cost data from 2010 cannot be meaningfully be added to cost data from 1998 without adjusting for inflation due to differences in purchasing power. The medical care component of the consumer price index is commonly used in economic evaluations to adjust for the effects of inflation on the cost of health care services. If the item of resource of interest is outside this sector of the economy, for example the wage cost for non-health sector workers, then this should be adjusted for using the relevant sub-scale in the index (i.e. wage inflation in the wider economy) (Lipscomb, Weinstein & Torrance, 1996; Corso & Haddix, 2003; Cairns & Fox-Rushby, 2005a; Culyer, 2005; Drummond et al, 2005; Glick et al, 2007; Meltzer & Smith, 2012).
There are three theories that are commonly used to explain why individuals (and societies) exhibit positive time preference: (i) the risk of the loss of future consumption opportunities; (ii) a preference for earlier rather than later consumption and; (iii) diminishing marginal utility over time. In the first of these theories, time preference is a consequence of a concern that death, or another event in the future such as a loss of income, may curtail future opportunities for consumption. In the second of these two theories, time preference reflects the preference that individuals have for earlier, as opposed to later, consumption and immediate utility as opposed to delayed utility. In the last of these theories, time preference is a consequence of the decline of marginal utility over time and a belief that future consumption will be higher, with the preference for immediate as opposed to later consumption (when the costs of two or more alternatives are equal) being a function of the value of the outcome of interest and the temporal delay until it can be realised. The third and last of these theories, that of diminishing marginal utility over time, is the one that is preferred by the majority of economists due to its simplicity and generality. Furthermore, within this theory, the decisions that individuals make in trade-offs between costs and outcomes that occur at different points in time, are not different in any significant way from the choices individuals normally make about consumption. There is simply a delay in consumption that needs to be taken into account when a choice is made about the desirability of the outcome. An important assumption of this theory is that of stationarity, that requires that an individual’s preference for one of two given outcomes is dependent on the period of time that separates them. The time preferences of individuals can be measured through their marginal intertemporal rate of substitution, the minimum sum required in the future to compensate them for delaying consumption. For example, if an individual is indifferent to whether they receive $1.03 one year in the future or $1.00 in the present day, their rate of time preference is 3 per cent per annum (Cairns, 2001; Frederick, Loewenstein & O’Donoghue, 2002; Cairns & Fox-Rushby, 2005a; Culyer, 2005; Berns, Laibson & Loewenstein, 2007; Glick et al, 2007; Kalenscher & Pennartz, 2008).
The time preference of individuals and societies is accounted for in economic evaluations through the use of discounting, a procedure that is used to convert future costs and outcomes to their present day values. To do this, future costs and benefits in an evaluation are re-expressed as their present day values through the use of an appropriate discount rate, which effectively places a declining weight on future events. The size of this weight is determined by the events distance into the future, with larger weights attached to events in the near future, and smaller weights attached to events in the distant future. The formula for the standard discounting model\(^{29}\) is \(PV = C / (1 + r)^t\) (time is in years) where \(PV\) = present value, \(C\) = cost; \(r\) = discount rate as a decimal fraction (i.e. 0.03) and \(t\) = time in years. A discount rate can also be expressed as a discount factor, which for year \(t\) and for an annual discount rate of \(r\), is calculated as \(1 / (1 + r)^t\). For a discount rate of \(r = 0.01\) for one year the discount factor is 0.909 and for five years it is 0.620. For example, in a comparison of two interventions with three year time horizons in which the annual costs are known: Intervention A, costs $5,000 in its first year, $10,000 in its second year and $15,000 in its third year and; Intervention B, costs $15,000 in its first year, $10,000 in its second year and $4,000 in its third year, it appears that the cost of intervention B is less than intervention A ($29,000 versus $30,000). When the two streams of future costs are converted to their present day value using a discount rate of 5% per annum, and it is assumed that costs are incurred at the end of each year, then the net present value of Intervention A is less than intervention B ($26,790 versus $26,810). This is the result of the majority of the cost of Intervention A being incurred in later years. In economic evaluations it is recommended practice to adjust for time preference if the period of follow-up is greater than one year. When discounting is not used in an economic evaluation where there are costs and outcomes that occur at different points in time, particularly in studies of interventions for chronic diseases and their prevention, then this can distort the relative Cost-effectiveness of an intervention and its alternatives and undermine the decision-making process (Cairns, 2001; Cairns & Fox-Rushby, 2005a; Culyer, 2005; Glick et al, 2007; Kalenscher & Pennartz, 2008; Meltzer & Smith, 2012).

There are two competing theories for the selection of an appropriate discount rate for use in an economic evaluation of a publically funded intervention, these are: (i) the social opportunity cost approach and; (ii) the social rate of time preference. The social opportunity cost approach has as its underlying premise that resources should not be committed to public projects if they produce a lower rate of return than private sector investments (the opportunity cost of capital). It is not in society’s interest to invest in a project that returns 2% per annum over 100 years, when it could put the same resources into an investment that returns 5% per annum over the same time period. For while the difference in the rate of return of the two investments is relatively small, over 100 years an investment that returns 2% per annum will return 18

\(^{29}\) Also known as constant rate discounting/exponential discounting (Cairns & Fox-Rushby, 2005a).
times less than one that returns 5% per annum. On this basis, the discount rate that should be used in an economic evaluation is the weighted average of the discount rates used in related private sector investments in the wider economy. The social rate of time preference approach is a reflection of the collective willingness of society to trade off consumption in the present day, against a higher level of consumption in the future. On this basis, the discount rate that should be used in an economic evaluation is the real rate of return produced by long-term government bonds. While strictly not an independent third theory, but a mix of approaches, the social price of capital approach is based on a premise that consumption is the underlying purpose of all economic activity (including investment) and consequently the true opportunity cost of an intervention is the present day value of the loss of the consumption needed to produce it. In this approach the costs and benefits of an intervention are transformed into a stream of consumption losses (the costs of an intervention, in terms of the foregone opportunities for consumption and investment), and a stream of consumption gains (the outcomes produced an intervention) which are then discounted using the social rate of time preference (Arrow et al, 1995; Lipscomb, Weinstein & Torrance, 1996; Cairns, 2001; Brouwer & van Exel, 2004; Cairns & Fox-Rushby, 2005a; Culyer, 2005; Drummond et al, 2005; Glick et al, 2007; Gravelle et al, 2007).

The theory that underlies the choice of an appropriate discount rate is complex and it is for this reason, and the need for standardisation for purposes of comparison, that guidelines have been developed in countries such as the United States that generally recommend that both costs and outcomes be discounted at 3% per annum. In practice the discount rate that is used in an economic evaluation is usually determined by the decision-making context, with the rate pre-specified by the decision-maker who is the target audience of the evaluation, or through convention in a given country. In most cases this falls between 3-5% per annum to reflect the consumption rate of interest (the time preference of individuals) and the rate of real long-term economic growth. Although, where possible an analyst should attempt to report the impact on the findings of the study using a range of discount rates, such as 5% per annum, so that meaningful comparisons can be made between the findings of different studies (Hunink et al, 2001; Brouwer & van Exel, 2004; Cairns & Fox-Rushby, 2005a; Drummond et al, 2005; Glick et al, 2007).
Incremental Cost-effectiveness Ratios

The results of Cost-effectiveness and Cost-utility analyses are commonly reported as Incremental Cost-effectiveness Ratios (ICERs), where the denominator is the difference in effectiveness (the net effect), and the numerator is the difference in cost (the net cost) of an intervention and a comparator. As previously discussed, in Cost-effectiveness and Cost-utility analysis there is no implicit decision rule for when an intervention should be implemented as these forms of economic evaluation (unlike Cost-utility analysis) only produce a net cost per unit of health outcome (i.e. the cost per life saved). In most cases a decision maker needs to decide whether the unit of health outcome has sufficient intrinsic value to justify the cost of achieving it. Interventions that are more effective than a comparator are usually also more costly, and ultimately a decision about whether or not they should be implemented is dependent on the resources available to the decision maker, the benefits that could be gained from the alternative use of those resources, and other considerations. It is only when an intervention dominates a comparator (it is less costly and more effective) or when the inverse is true, and an intervention is dominated by a comparator (it is more costly and less effective) that a decision can be easily made on the basis of efficiency. In most cases the Cost-effectiveness or cost-utility of an intervention can only be determined in relation to other interventions that society has chosen to implement, such as dialysis for end-stage renal disease, or in relation to a known maximum acceptable ratio of costs to outcomes, such as $50,000 AU a QALY (Garber et al, 1996; Gold et al, 1996; Mandelblatt et al, 1996; Seigel, Weinstein & Torrance, 1996; George, Harris & Mitchell, 2001; Drummond et al, 2005; Fenwick & Byford, 2005; Briggs, Claxton & Schulpher, 2006; Briggs et al, 2012).

30 Alternatives can also be excluded through extended dominance – where a linear combination, ie. 50% of patients receive one of the alternatives and 50% receive another of the two alternatives, is less effective and more costly than a comparator (Gold et al, 1996; Siegel, Weinstein & Torrance, 1996; Cairns & Fox-Rushby, 2005b; Culyer, 2005; Drummond et al, 2005).
The results of an economic evaluation can be represented on a Cost-effectiveness plane, a diagram with separate quadrants that are mapped to the points of a compass, with the difference in effect displayed on the horizontal axis and the difference in cost displayed on the vertical axis. The maximum threshold for the amount a decision-maker is willing to pay for a unit of health outcome (the highest acceptable Cost-effectiveness ratio, i.e. the maximum cost per QALY) is represented by the slope of a straight line that passes through the origin (refer to Figure 1). If the slope of the line that represents the incremental Cost-effectiveness ratio of the intervention is less than a decision-makers acceptable threshold then, all else being equal, the intervention should be implemented. Interventions that are dominant are located in the SE quadrant of the plane and interventions that are dominated are located in the NW quadrant. Interventions that are more effective and more costly than the comparator are located in the NE quadrant and interventions that are less effective, while also being less costly, are located in SW quadrant (Drummond et al, 2005; Fenwick & Byford, 2005; Briggs, Claxton & Schuler, 2006).

**Figure 1: Cost-effectiveness plane**

Source: Petrou & Gray (2011)
Uncertainty and sensitivity analysis

There are four broad areas of uncertainty in economic evaluations that are the result of: (i) the decisions that are made during the process of framing the analysis; (ii) the data requirements of the analysis; (iii) the extrapolation of the results of primary studies and; (iv) the generalisability of the results of the analysis. The first of these areas of uncertainty is a result of disagreements among health economists about analytic issues such as the choice of study perspective and the methods that should be used for the measurement and valuation of costs and outcomes. The second is the result of the uncertainty produced by natural variations in populations and how these impact on resource usage and healthcare outcomes. The third relates to the use of modelling to extend the time horizon of an evaluation beyond the period of primary data collection through the use of data obtained from other studies, and assumptions relating to issues such as the persistence of the effect of an intervention. The fourth relates to the degree to which the findings of an economic evaluation conducted in a specific context and target population can be generalised to other contexts and populations (Briggs & Gray, 1999; Briggs, 2001; Drummond et al, 2005).

For the findings of an economic evaluation to be useful to a decision-maker there needs to be an explicit examination of uncertainty and the impact that it has on the results of the analysis. There are three main types of uncertainty and sensitivity analysis: (i) simple one way/univariate analysis; (ii) multi-way/multivariate analysis; and (iii) probabilistic sensitivity analysis. In a simple one way/univariate sensitivity analysis, one or more of the variables about which there is uncertainty are varied across a plausible range, while all other variables are held constant. While this form of sensitivity analysis provides a useful means of uncovering some of the key drivers of the results of an evaluation, it provides only an incomplete picture of true extent of the overall uncertainty surrounding the results. The reasons for this are threefold; (i) the incremental cost and effectiveness and cost of an intervention are the product of more than one variable; (ii) interactions between variables can produce results that are substantially different from the sum of the individual variables and; (iii) the uncertainty for a ratio of two uncertain numbers can be greater that the uncertainty for the two numbers in isolation. In a multi-way/multivariate sensitivity analysis, two or more variables are varied simultaneously to examine their combined impact. It can be difficult to report the results of a multi-way/multivariate analysis when the combined effects of a large number of variables are examined simultaneously. One solution to this problem is the use of scenario and extreme scenario analysis. In scenario analysis, the values of the components about which there is uncertainty are used to represent a series of different, but plausible, states of the world. In extreme scenario analysis, a best case and a worst case scenario is constructed based on the most optimistic and most pessimistic values for the components of an evaluation. If an intervention is shown to be cost-effective in a base case analysis and this is supported by the results of an extreme scenario analysis, the
findings of the evaluation can be treated with a high degree of confidence. Extreme scenario analysis is an efficient method of dealing with uncertainty in situations where the baseline value and range of a variable is known, but not its distribution. It is less useful in dealing with uncertainty that is the result of the analytic methods that were used in an evaluation that cannot be easily characterised as either optimistic or pessimistic (Briggs & Gray, 1999; Briggs, 2001; Shaw & Zachry, 2002; Walker & Miners, 2005; Andronis, Barton & Bryan, 2009; Jain, Grabner & Onukwugha, 2011; Meckley et al, 2010; Eddy et al, 2012).

Probabilistic sensitivity analysis was developed to address the shortcomings of univariate and multivariate sensitivity analysis. It does this by simultaneously examining the joint uncertainty of all of the components in an evaluation through the use of Monte Carlo simulation, in which the components are randomly sampled across their probable ranges and distributions. The simulation is run a large number of times (i.e. 1000 to 10,000 times) to produce an estimate of the distribution of the result, and an associated measure of variance. The distributions that are attached to the variables in a probabilistic sensitivity analysis are not arbitrary, with the selection of the choice of the distribution guided by the type of variable, the form of the data and the process that was used in its estimation. In general these distributions are the same as those that are commonly found in textbooks of statistical methods for the analysis of clinical research data (i.e. normal, binomial etc.). The results of a probabilistic sensitivity analysis are commonly reported in graphical form, as a scatter plot of the joint distribution of incremental costs and effects on a Cost-effectiveness plane. The scatter plot (which typically covers all of the four quadrants of the Cost-effectiveness plane) summarises the uncertainty surrounding an ICER by illustrating the probability that an intervention is more cost-effective than an alternative for any given value that a decision maker is willing to pay for an additional unit of health. This is demonstrated by the proportion of points on the scatter plot that fall in the south-east quadrant of the Cost-effectiveness plane, beneath a line that represents the decision maker’s maximum acceptable ratio of costs to benefits. It is not unusual for an economic evaluation to use univariate and multivariate sensitivity analysis to examine the impact of uncertainty that is a result of decisions made about the structure of a decision-analytic model, and probabilistic sensitivity analysis to examine the impact of parameter and overall uncertainty (Briggs & Gray, 1999; Briggs, 2005; Claxton at el, 2005; Drummond et al, 2005; Briggs, Claxton & Schulpfer, 2006; Vergel & Schulpfer, 2006; Caro et al, 2012).
There are several methods that can be used to quantify and present the uncertainty surrounding the results of an economic evaluation, the first of these being the calculation of confidence intervals for the incremental cost-effectiveness ratios. This can be done using parametric approaches such as Fieller’s theorem (Fieller, 1954), or non-parametric approaches such as bootstrapping. The method developed by Fieller dates back to work undertaken in the 1930s on ratio statistics. The method assumes that the differences in cost and effect in the ICER are normally distributed and have a known variance. The confidence interval calculated from the difference in cost and its variance, the difference in effect and its variance, and the covariance between costs and effects. The usefulness of this method is limited by its underlying assumption of normality, as in practice sampled cost data is usually highly skewed.

Bootstrapping is a computer based technique in which a large number of random samples, of the same size as original sample, are taken from the observed sample of the costs and outcomes that were used to calculate the ICER, from which an empirical estimate of the sampling distribution of the ICER is generated. The approach is based on two dependent assumptions: (i) as the sample size increases its distribution moves closer to that of the population distribution and; (ii) if the first assumption is true, then as the number of replications increases the sampling distribution will move closer to the true sampling distribution. While there are no formal rules that specify the number of bootstrapped replications that are required to produce a reliable result, in general it is recommended that approximately 1000 replications be undertaken. Unfortunately, an ICER is a ratio-based statistic with an unstable distribution that is derived from two random variables (incremental cost and incremental effect) both of which can be zero, and consequently in circumstances where there is no difference in incremental effect or incremental cost, it can be very difficult (if not impossible) to calculate a confidence interval for an ICER using statistical techniques such as those listed above (Briggs, 2001; Fenwick, Claxton & Schulpher, 2001; Shaw & Zachry, 2002; Fenwick, O’Brien & Briggs, 2004; Briggs, 2005; Drummond et al, 2005; Fenwick & Byford, 2005; Briggs, Claxton & Schulpher, 2006; Briggs, Wonderling & Mooney, 1997; Wang & Zhao, 2008; Jain, Grabner & Onukwugha, 2011; Meckley et al, 2010).

The difficulties outlined above, and other issues encountered in the calculation of confidence intervals for incremental cost-effectiveness ratios led to the development of Cost-effectiveness acceptability curves (CEACs) and the net-benefit framework, as relatively simple alternatives for the presentation of uncertainty in economic evaluations. A cost-effectiveness acceptability curve is a plot of the proportion of the incremental cost-effectiveness ratios (in a probabilistic sensitivity analysis) that are cost-effective for a given maximum acceptable ratio of costs to benefits. The net-benefit approach is based on simple re-arrangement of Cost-effectiveness decision rule (incremental benefits divided by incremental costs must be less than a maximum acceptable ratio) that places the costs and outcomes of an intervention on a single scale of either net health benefit, or net monetary benefit (the most widely used of the two). This
produces a linear expression, which for large sample sizes has a normal distribution, which makes it possible to use standard statistical techniques for the quantification of uncertainty. A net health benefit is deemed to be positive if the gain in health produced by an intervention is greater than that produced by the use of the same resources for a marginally cost-effective intervention. A net monetary benefit is achieved and an intervention is deemed to be cost-effective, when the increase in effectiveness multiplied by the maximum amount a decision-maker is willing to pay for a unit of health outcome is less than the difference in cost (Stinnett & Mullahy, 1998; Fenwick, Claxton & Schulpheer, 2001; O'Brien & Briggs, 2002; Briggs, O'Brien & Blackhouse, 2002; Shaw & Zachry, 2002; Fenwick, O'Brien & Briggs, 2004; Briggs, 2005; Cuyler, 2005; Drummond et al, 2005; Fenwick & Byford, 2005; Briggs, Claxton & Schulpheer, 2006; Briggs, Wonderling & Mooney, 1997; Meckley et al, 2010; Jain, Grabner & Onukwugha, 2011; Petrou & Gray, 2011).
Conclusion

Welfare economics is an analytical framework for assessing the desirability of the alternative allocation of resources for the production and distribution of goods and services. The relative desirability of these alternatives is determined from the sum of individual utility, which is maximised as a result of the choices made by individuals in a free and perfectly competitive market. A state of competitive equilibrium exists within this market as a result of consumers acting to maximise their utility, and producers acting to maximise their profit. If there are no increasing returns from production, and every good or service that an individual may wish to consume is priced and available in the market, then this will result in an allocation of resources that is Pareto optimal for any given distribution of purchasing power. There is consequently no other allocation of resources that would improve the utility of one individual, without decreasing the utility of another. It is generally accepted that the characteristics of healthcare as a commodity deviate significantly from those required for a perfectly competitive market, and that this has resulted in pervasive market failure and the sub-optimal and inefficient allocation and use of resources. As a consequence of this failure, non-market mechanisms such as economic evaluation were developed to provide a method for the explicit evaluation of the costs and benefits of alternative healthcare interventions.

As an aid to decision-making economic evaluation can only be used where the objective of an intervention is clearly articulated, and there is at least one alternative course of action against which an assessment of efficiency can be made. The choice of the form of evaluation to use for a given analysis (CBA, CEA or CUA) is determined by the question that needs to be answered, the information needs of the decision maker, and the selection of outcome measures and the availability of data. Unlike CBA, CEA and CUA have no implicit decision rule for when an intervention should be implemented, as they only return a net cost per unit of health outcome. Consequently it is left to decision makers to judge whether the unit of health outcome has sufficient intrinsic value to justify the cost of achieving it. This can be relatively straightforward when there is a clearly divisible set of interventions, a fixed budget, and the maximisation of health is the only objective. In situations where the budget is not fixed, and the implementation of an intervention may require the re-allocation of resources from other sectors of the economy, all a decision-maker can do is set a threshold value for the ratio of costs to benefits that must be achieved before an intervention can be considered for implementation.

In the next chapter of the thesis, the reader is provided with an introduction to the use of decision analytic modelling in the economic evaluation of healthcare interventions.
Chapter 4: Decision analytic modeling

Introduction

In the preceding chapters of the thesis the reader was provided with an introduction to DMD and the alternative treatment pathways for the use of LTMV, the theory that underlies the explicit economic evaluation of healthcare interventions and the methods that are used in Cost-effectiveness and Cost-utility analysis. The aim of this chapter is to provide the reader with an introduction to the use of decision analytic modelling in economic evaluations and how a model is constructed, analysed and validated.

The chapter starts with an introduction to decision analytic modelling and how it is used to structure imperfect information in situations where there is significant uncertainty, multiple competing objectives and a clear need to determine the option that returns the highest expected value. This is followed by a discussion of how a decision about the type and structure of a model for use in an economic evaluation is primarily determined by the characteristics of the disease and intervention under evaluation. This is in turn followed by a detailed examination of decision analytic and Markov modelling and an overview of some newer modelling techniques, specifically micro-simulation (individual level/patient level) modelling, discrete event simulation modelling and dynamic transmission modelling. The construction and analysis of a decision analytic model is then described in detail, as is the construction and analysis of a Markov model. The chapter ends with the methods that are used to validate decision analytic models.
Decision analytic modeling

Decision analytic modelling is a systematic framework that is used to structure imperfect information about events and outcomes in situations where there is significant uncertainty, multiple competing objectives and a clear need to determine the option that returns the highest expected value. Its use in economic evaluation is underpinned by expected utility theory. It is a process through which the expected value of at least two alternatives is determined from the probability of their occurrence, and the sum of their respective costs and outcomes. In the economic evaluations of healthcare interventions, decision analytic modelling is used to synthesize the best available evidence about the incidence and natural history of a disease, the effectiveness of alternative treatment options and their impact on outcomes such as survival and quality of life. It is an approach that is used in a variety of situations, for example where: (i) a clinical trial of an intervention is yet to be conducted; (ii) clinical trials have been conducted but there was no data collected about the use of healthcare resources; (iii) the outcome measures of the available clinical trials address only short term intermediate outcomes such as a reduction in blood pressure and not important long-term patient relevant outcomes such as a reduction in the risk of cardiovascular disease; (iv) there is substantial uncertainty surrounding the best available evidence for a complex time critical clinical or policy decision; (v) a clinical trial cannot be conducted due to the cost of doing so, or political or cultural considerations, time pressures or ethical constraints (Buxton et al, 1997; Halpern, McKenna & Hutton, 1998; Karnon & Brown, 1998; Briggs, 2001; Kuntz & Weinstein, 2001; Goldie & Corso, 2003; Drummond et al, 2005; Briggs, Claxton & Schulpher, 2006; Gray et al, 2011; Caro et al, 2012).

Model structure and type

The choice of the type of model to use in an economic evaluation and its structure, is driven by the characteristics of the disease and intervention under evaluation, characteristics such as: (i) the nature of the disease, i.e. chronic as opposed to acute; (ii) the number of possible health states; (iii) the nature of the risk of future health-related events and whether these change over time; (iv) the effectiveness of treatment and whether this remains constant over time; and (v) whether or not the risk of the occurrence of a future health related event is determined by the previous health related events experienced by a given individual. Ideally, a model should reflect the key features of the natural history of the disease of interest, with the events in the model corresponding to biological or clinically important health states. Models are at best only simplified representations of reality and while there are no absolute rules that can be used as the basis for a decision about their structure and complexity, ideally they should be no more complex than is absolutely necessary for them to be able to provide a useful basis for decision-making. A complex model may be better able to incorporate all of the aspects of a disease and its treatment that content experts may
claim is necessary but it may be next to impossible to understand and to communicate and be extremely
difficult to construct, debug and evaluate. In practice, the structure and complexity of a model is usually
determined by the availability of the data needed for the description of the model parameters (Mandelblatt
et al, 1996; Karnon & Brown, 1998; Goldie & Corso, 2003; Lang, Lopert & Hill, 2003; Drummond et al,
2005; Brennan, Chick & Davies, 2006; Briggs, Claxton & Schulpher, 2006; Gray et al, 2011; Caro et al,
2012; Roberts et al, 2012).

Decision-analytic models can be longitudinal, cross sectional, deterministic or stochastic. In a longitudinal
decision-analytic model a homogenous cohort is followed over time to estimate the long-term outcomes
(i.e. quality adjusted life-expectancy) that are produced by alternative treatment pathways or interventions.
This form of decision analytic modelling is widely used in economic evaluations to extend the time horizon
of an analysis to beyond the period of primary data collection, (i.e. to the life-time of the study
participants). In a cross sectional decision analytic model, the health status of a defined population group
(i.e. Australian women aged 50-75) is measured at different points in time to calculate outcome such as
quality adjusted life years. Deterministic decision-analytic models are constructed using data about which
there is a reasonable degree of certainty, for example, the mean number of cases of a disease that are
expected to develop in a single year. Stochastic decision-analytic models are used in situations where
there is significant uncertainty about the number of disease related events in a given period of time. The
construction of a stochastic decision analytic model generally requires much more data than a
deterministic model. It also requires a detailed understanding of the disease process and the intervention,
the relationships between events and how this impacts on model parameters, for example whether two or
more events are independent of each other, or whether they interact to increase or decrease the
probability of future events (Mandelblatt et al, 1996; Last , 2001).

Decision trees are the simplest and most widely used form of decision analytic modelling. They are of
most use when the scenario they are used to represent is relatively uncomplicated and there is a limited
number of chance events and little or no reoccurrence of disease states. In general decision analytic
models are used to represent disease processes, or events, which occur over relatively short periods of
time, as the representation of time dependent events can be difficult to implement. If used to represent
chronic diseases with a large number of possible health states, or reoccurring events, then decision
analytic models can become extremely large and overly complicated. This is because every health state
and event in the model needs to be represented as a separate branch of the tree. These diseases are
better represented through the use of some form of state transition model such as a Markov model, in
which a disease is broken down into a series of health-states, such as treatment, remission, progression
and death which individuals transition between over fixed periods of time. It is not unusual for a model to
contain decision analytic and Markov components, for example where a decision tree is used to represent the events and activities of an intensive but relatively short-term diabetes prevention program (i.e. less than 1 year duration) and a set of Markov nodes are used to model the lifetime costs and outcomes (Mandelblatt et al, 1996; Briggs & Schulpher, 1998; Kuntz & Weinstein, 2001; Goldie & Corso, 2003; Briggs, Claxton & Schulpher, 2006; Sun & Faunce, 2007; Gray et al, 2011; Petrou & Gray, 2011).

There are two broad types of Markov models that are distinguishable by form of transition probabilities that are used. The first and most commonly used type are Markov process models that have time-dependent transition probabilities, where the probability of moving between health-states changes over time. This feature makes this form of Markov modelling well suited for use in economic evaluations of interventions for chronic diseases, where the risk of morbidity and mortality increases with age. The second and less commonly used type are Markov chain models that have fixed transition probabilities, where the probability of moving between health states does not change over time. This feature limits the use of this type of Markov modelling to economic evaluations of healthcare interventions for diseases with short time horizons, in which increasing age; changes in disease severity do not need to be taken into account. An important limitation of both of these forms of modelling is what is known as the Markovian assumption. In a Markov model it is assumed that: (i) all of the individuals in a given health state are indistinguishable from each other; (ii) there is no memory of the previous health states that an individual may have resided in, or of the disease events and treatments they may have received and (iii); the probability of a transition between a health state is determined only by the health state that an individual is currently residing in. This means that, for example, in a Markov model of the treatment of myocardial infarction, for a given health state, there is no ability to take into take into account individual variations in disease history and treatment, such as the number of previous myocardial infarctions, the treatment received, or their age related risk of death. While a Markov model can be designed to take into account the past events in the history of a disease, for example through the use of one or more complex health states with time dependent transition probabilities, this is not a straightforward process and one that could result in a model that is overly complex and extremely difficult to understand (Sonnenberg & Beck, 1993; Briggs & Schulpher, 1998; Petitti, 2000; Kuntz & Weinstein, 2001; Briggs, Claxton & Schulpher, 2006; Sun & Faunce, 2007; Petrou & Gray, 2011).
Micro-simulation modeling and other alternative approaches

In the last ten years there have been a range of new modelling techniques that have begun to be used in economic evaluations, techniques such as micro-simulation (individual level/patient level) modelling, discrete event simulation modelling and dynamic transmission modelling. A micro-simulation model, unlike a Markov model, is not memory-less and does not require that events occur within pre-determined and fixed cycles of time. Individuals move through the model one at a time and their history determines the transition probabilities of their future events and the costs and health outcomes they accumulate. In a discrete event simulation model, as in a micro simulation model, events can occur at any point in time and an individual's history can be taken into account. There is also the ability to take into account interactions between individuals. For example, discrete simulation modelling can be used to represent the activities of a renal transplant unit where a decision to operate on one individual of a group awaiting a transplant impacts on the costs and outcomes of the remaining individuals. Individuals within a discrete simulation model can also be assigned attributes such as age, stage of cancer etc., that influence their pathway through the model and they can acquire new attributes as they move through the model as a result of the treatments they receive, and other events. Consequently, the pathway that an individual follows through a discrete event simulation model, and their accumulated costs and outcomes, are a product of the attributes they enter the model with, the attributes they gain while moving through the model, the events they experience and the interactions they have with other individuals. Dynamic transmission models are increasingly being used in economic evaluations of interventions for infectious diseases where changes in factors over time, such as the rate of vaccination indirectly modify the risk of disease for susceptible individuals. The impact of indirect effects be can difficult to represent in decision analytic models, but they are easily accounted for in dynamic transmission models through the use of feedback loops and time delays (Karnon & Brown, 1998; Briggs, Claxton & Schulpher, 2006; Grey et al, 2011; Petrou & Gray, 2011; Caro et al, 2012; Karnon et al, 2012; Pitman et al, 2012).

The design, construction and analysis of a decision-analytic model

The design, construction and analysis of a decision analytic model is a five step process. In the first step of the process the nature and scope of the question is identified and systematically broken down into its consistent components. As in the specification of a study question in an economic evaluation the objective of the analysis is defined, along with the target population, the alternatives under evaluation and the outcomes of interest. A decision is then made about the boundaries of the evaluation that takes into consideration, the complexity of the problem, the needs of the decision-maker and the availability of data. In the second step of the process a decision tree is developed that graphically depicts the components of
the decision problem and its consequences. The “tree” is constructed working from left to right starting with the alternative treatment options, which branch off a square decision node, the first point of choice in the model. These branches attach to circular chance nodes which are used to represent two or more possible alternative events that an individual may or may not experience, events which themselves may lead to further chance nodes that correspond to subsequent events about which there is uncertainty and so on. The branches at these chance nodes are labelled with the events they represent, i.e. disease reoccurrence, remission or death. These events need to be mutually exclusive and exhaustive, with the sum total of the probability of their occurrence being exactly equal to one. The last component of the tree is a triangular terminal node that represents the final event in the pathway, such as recovery, cure, disability or death (Mandelblatt et al, 1996; Tom & Schulman, 1997; Karnon & Brown, 1998; Petitti, 2000; Lang, Lopert & Hill, 2003; Goldie & Corso, 2003; Drummond et al, 2005; Briggs, Claxton & Schulpher, 2006; Gray et al, 2011; Caro et al, 2012).

In the third step of the process the probability and payoffs (cost and outcomes) that are needed to complete the model are identified, collected and attached to the appropriate points in the decision tree. By convention the probability of an event at a chance node is placed to the right of the node, underneath the label for the event that it represents. These probabilities range from 0.00 (complete certainty that the event will not occur) to 1.0 (complete certainty that the event will occur). The probabilities in a decision analytic model can be: (i) joint – two events that occur at the same time; (ii) conditional – the occurrence of an event, given that another event has occurred and; (ii) independent – two or more events that occur independently of each other. The probability that is attached to the left most chance node in a decision analytic model is indicative of the likelihood of that event happening, with the probabilities in the chance nodes to its right being conditional on the occurrence or non-occurrence of the previous event. The payoffs for the costs and final outcomes (i.e. utilities, life-years gained, quality adjusted life-years etc.) for the sum of the events for each of the alternative pathways through the decision tree are then attached to the terminal nodes. Unlike the detailed and well established methods for locating and assessing the evidence for the effectiveness of an intervention for use in an economic evaluation, relatively little has been written about the methods that should be used to collect probability, cost and other data that are needed to populate a model. It is therefore important that the methods used to do this are systematic and clearly documented, so that the end users of the model are able to form a judgement about the appropriateness and quality of the data that was used in its construction. In the fourth step of the process the model is validated. In the fifth and final step of the process the tree is analysed by the summing the costs for the event pathways for each of the alternatives, by the weight of the probability of their occurrence and their expected values (outcomes). A tree can also be “rolled back”. A process where,

31 The methods that are used to do this are described later in the chapter.
working from right to left, the expected values (outcomes) and costs for each of the terminal nodes are multiplied by the probability of the likelihood of their occurrence, which are then summed for each of the preceding chance nodes (Mandelblatt et al, 1996; Tom & Schulman, 1997; Karnon & Brown, 1998; Lang, Lopert & Hill, 2003; Drummond et al, 2005; Brennan, Chick & Davies, 2006; Briggs, Claxton & Schulpher, 2006; Gray et al, 2011; Caro et al, 2012).

The next figure, Figure 2, contains an example of a simple decision analytic model that compares a policy of bi-annual screening for breast cancer to that of no screening. The probability of breast cancer being present and the sensitivity and specificity of the screening tests are defined as variables in the model's root node. For each approach there are two possible pathways; (i) that breast cancer is present and; (ii) that it is not present. In the no screening arm of the model there are two possible final outcomes; (i) cancer, when breast cancer is present and; (ii) when it is not present. In the screening arm of the model there are four possible outcomes: (i) that cancer is present and it is detected by the screening test (true positive); (ii) that cancer is present but it is not detected by the screening test (false positive); (iii) that cancer is not present but the screening test is positive (false positive) and; (iv) that cancer is not present and the screening test is negative (true negative).

**Figure 2: Example of a decision analytic model**

![Decision Analytic Model](image)

Source: Gray et al (2011, pp. 91-194)
The design, construction and analysis of a Markov model

As previously described, in a Markov model the natural history of a disease is broken down into a finite set of mutually exclusive and collectively exhaustive health states that are used to represent important events such as treatment, remission, reoccurrence and death. The first step in the eight step process of the design, construction and analysis of a Markov model is the identification of these health states, and the allowable transitions from one state to another (i.e. treatment to remission, treatment to death, reoccurrence to treatment, reoccurrence to death). The health states in a Markov model can take one of the following forms: (i) transient, a state of health that can be revisited at any time; (ii) temporary, an important health state (i.e. a cycle of chemotherapy) that lasts for only a short period of time (i.e. a month) after a transition must be made to another health state; (iii) tunnel, a sequence of events that lasts for more than one cycle that must be followed in a fixed order and; (iv) absorbing, a health state such death, that once entered cannot be left (i.e. there are no possible transitions to any other health state). The number of health states in a Markov model should ideally be kept to a small a number as possible for as their number increases so does the complexity of a model and the amount of data needed to populate it (Sonnenberg & Beck, 1993; Briggs & Schulpfer, 1998; Fox-Rushby & Fidan, 2005; Claxton & Schulpfer, 2006; Sun & Faunce, 2007; Gray et al, 2011; Petrou & Gray, 2011; Siebert et al, 2012).

The second step in the eight step process is the determination of the length of the Markov cycle. This is the minimum period of time (of equal length) i.e. hours, weeks, months or years that can be spent in a health state before a transition can be made to another state. The choice of cycle length should ideally reflect the natural history of the disease being modelled or a biological process specific to that disease and the frequency of important events such as treatment, reoccurrence or remission. The third step of the process is the determination of initial and transition probabilities. The first of these two sets of probabilities is used to define the initial distribution of the cohort being modelled across the health states of the model prior to its first cycle. For example, in a model with three health states, i.e. healthy, unwell and dead in which everybody is alive at the start of the model but 20% have the disease of interest, then the initial probabilities would be; healthy 0.80; unwell 0.20 and; dead 0.00. The transition probabilities define the movement between health states for each cycle of the model. These probabilities can be constant over time or time dependent and their sum total must be exactly equal to 1.0. In the fourth step of process rewards are assigned to the health states. There are different types of rewards, for example: state rewards, which are costs and outcomes (i.e. utilities, life-years) that are associated with spending one cycle in a health state and; transition rewards, which are non-recurrent costs or outcomes that are accrued only when there is a transition into a health state (Fox-Rushby & Cairns, 2005; Fox-Rushby & Fidan, 2005; Gray et al, 2011; Siebert et al, 2012).
In the fifth step of the process a series of decisions are made about the need for a stopping rule, the need for discounting and the need for a half cycle correction. A stopping rule is required for models in which there is no absorbing health state such as death (as in the model in Figure 3) that can bring the process to an end. A stopping rule is also required in situations where a point of equilibrium is reached after a number of cycles and only a small proportion of the cohort continues to move between the health states in the model, or where as a consequence of the structure of a model less than 100% of the cohort will reach an absorbing state within a given numbers of cycles (i.e. 20 years). In such cases, the model can be forced to stop when the transitions between health states falls below a specified level or a specified number of cycles (i.e. 40 six month cycles, or 20 years) (Gray et al, 2011).

For a model with a time horizon of longer than one year, discounting will need to be applied to all costs and outcomes. For a model with a Markov cycle length of one year, the cycle number (in years) is used as an input into the standard discounting formula. In a Markov model it is assumed that the transitions between health states all take place at the start of the Markov cycle, and that the members of the cohort who transition into an absorbing health state, such as death, do this immediately at the start of the cycle. In the everyday world individuals move through a disease process and its treatment on a continuous basis with events such as death occurring at any time. To assume that these events occur at the start of a Markov cycle can lead to the miscalculation (i.e. underestimation/over estimation) of the proportion of the cohort who are resident in health states such as alive or dead and outcomes such as survival and life-expectancy. The potential for bias that is the result of this assumption can be corrected through a half-cycle correction, which treats these events as if they occur, on average, at the midpoint of a cycle. There are two methods that are used for a half cycle correction: (i) the state membership at time \( t \) can be added to the state membership of time \( t+1 \), which is then multiplied by the cost and outcome associated with that health state and then divided by 2 or; (ii) 0.5 is added to the estimate of total life expectancy or a half a cycles worth of incremental cost or utility is added to the cumulative totals for each state in the model. While a half-cycle correction is recommended practice, it is unlikely to have a major impact on the results of an analysis due to the comparative nature of economic evaluation and its focus on the magnitude of the incremental difference in the costs and outcomes of an intervention above those of a comparator. It is only in circumstances where the time horizon of a model is long and the cycle length as a proportion of the time horizon is long that the lack of a half cycle correction could be significant problem (Briggs & Schulpfer, 1998; Briggs, Claxton & Schulpfer, 2006; Gray et al, 2011; Siebert et al, 2012).
In the sixth step of the process of the development and analysis of a Markov model, a decision is made about the form of analysis that will be used. In the seventh step of the process, the model is validated (details in the next section) and in the eighth and final step of the process the model is evaluated. There are two methods that are commonly used to analyse a Markov model: (i) cohort analysis and; (ii) first order Monte Carlo simulation. In a cohort analysis the individuals in the model usually start in the same initial health state, the simulation is run and the cohort is followed for the duration of the model. During each cycle of the simulation the individuals move between health-states as determined by the transition probabilities. A profile is generated of the number of patients in these health-states for each cycle of the simulation, as well as the total number of cycles spent in each health state and their accumulated costs and outcomes. The main advantage of this approach is that it is a relatively easy to use and transparent method of evaluation that is understood and accepted by policy makers. Its main disadvantage, a consequence of the Markovian assumption is the requirement that all of the relevant clinical information for the members of the cohort must be represented in the health states in the model. This can result in the need for a large number of health states and extremely large and complex models. In a first order Monte Carlo simulation a cohort of individuals move through a model, one individual at a time, while a profile of their disease pathways, costs and outcomes are generated. The simulation begins with the individuals in an initial health-state, with the state transition probabilities and a random number generator used to determine the health state that they will begin the next cycle in. This model continues to cycle until each of the individuals reaches an absorbing state, such as death, and the simulation ends. The main advantage of this approach is that it generates a mean and an associated estimate of variance as the path followed through the model by an individual is subject to random variation. The main disadvantage of this approach is that it is a less transparent method of evaluation than cohort simulation, as it only produces a profile of the accumulated costs and outcomes for each individual, as opposed to a profile of the outcomes and costs for each cycle of the model (Mandelblatt et al, 1996; Sonnenberg & Beck; 1993; Tom & Schulman, 1997; Briggs & Schulpher, 1998; McCabe & Dixon, 2000; Kuntz & Weinstein, 2001; Weinstein et al, 2003; Fox-Rushby & Digan, 2005; Briggs & Schulpher, 2006; Sun & Faunce, 2007; Gray et al, 2011).
An example of a Markov model can be found in Figure 3. This model was used by Iskedijian et al (2005) in an economic evaluation of the use of Avonex (Interferon beta-1a) as an early treatment for Multiple sclerosis, a chronic progressive neurological disease. The evaluation examined the costs and outcomes produced by Avonex and whether its initiation after a single demyelinating event (SDE) could delay the development of a second demyelinating event and as a consequence of this clinically definite Multiple sclerosis (CDMS). The impact of Avonex on the progression of the disease, post CDMS, was measured through the use of the Expanded Disability Status Scale (EDSS) and represented in the model in terms of increasing severity of disease (EDSS 1 through to EDSS (level) 6+). Individuals enter the model after a single demyelinating event and are treated with Avonex. If there no further events they remain monosymptomatic (at the top right of the model). If they have another demyelinating event and meet the criteria for CDMS they then move to the node (down and to the left) that corresponds to the level of their disability (EDSS level 1 to 6+). For each cycle of the model, if the severity of their disease is unchanged they remain in the same health state (i.e. EDSS level 1 to EDSS level 1). If their disease increases in severity they can transition down the model to another health state with a higher level of disability (i.e. from EDSS level 1 to EDSS level 3). If the disease decreases in severity they can move back up the model (i.e. from EDSS level 3 to EDSS level 2 or EDSS level 1) but they cannot move back to the monosymptomatic state or the CMDS health state.
Figure 3: Example of a Markov model

Model validation

For a model to be useful to a decision maker they need to be confident that it accurately reflects the impact that an intervention has on an outcome of interest. A modeller can promote confidence in their model and in the credibility of the results it produces through transparency and careful validation. For a model to be transparent there needs to be a careful, detailed, non-quantitative description of the model's structure, parameters, equations and underlying assumptions. In addition to this overview, there should be a detailed technical description of the model and the data that was used to construct it, at a level of detail that would allow an end user to replicate it, if they wished to do so. The aim of the non-quantitative overview of the model and the detailed technical description is to provide the end user with information that they can use to understand what the model does, and how it does it, its limitations, potential accuracy and whether or not it is suitable for their needs. The validity of a model is how well it does what it was intended to do and the accuracy of the outcomes it produces. In most cases models are developed in situations where there is substantial uncertainty and little or no directly observable data available about the outcomes of interest. In most cases this means that models cannot be directly validated, for if the data to do this was readily available then there would be little or need for a model (Mandelblatt et al, 1996; Seigel, Weinstein & Torrance, 1996; McCabe & Dixon, 2000; Weinstein et al, 2003; Fox-Rushby & Digan, 2005; Eddy et al, 2012).

There are four main forms of model validity: (i) descriptive validity; (ii) internal validity; (iii) external validity and; (iv) predictive validity. The descriptive validity of a model is the extent to which its structure and underlying assumptions are reasonable and consistent with what is known about the natural history of the disease being modelled and its clinical management. There are four characteristics that are particularly important in determining the descriptive validity of a model: (i) structure – the extent to which the model incorporates all of the components of the disease and the intervention that are important to context experts and whether the relationships between these components is an accurate reflection of the current state of medical knowledge; (ii) data sources – their quality and the methods that were used to obtain them (i.e. a formal systematic review as opposed to expert option); (iii) problem formulation – the intervention and outcomes that were examined in the model, the setting, target population, time horizon and any simplifying assumptions and; (iv) results – the degree to which the outcomes produced by the model meet the expectations of context experts (Mandelblatt et al, 1996; Weinstein, et al, 2001; Gray et al, 2011; Petrou & Gray, 2011; Eddy et al, 2012; Roberts et al, 2012).
As previously discussed, decision analytic models are at best only simplified representations of reality that should ideally be no more complex than is strictly necessary to address the problem at hand. The descriptive validity of a model cannot be established through a simple application of a set of unambiguous criteria that must be met for it to be seen to be “valid”. It is not a fixed property of the model but something that can vary as a consequence of the setting and context of the particular application that it is used for. While an analyst may claim that their model has descriptive validity, stakeholders such as clinicians or patient groups may dispute this if they are unhappy with the results it produces. These groups may, for example, claim that a model is too simplistic and that it fails to reflect the complexity of clinical practice, or the everyday experiences of patients. While this may be a valid criticism, a complex model of an intervention cannot be constructed unless the data needed to do this is readily available. The descriptive validity of a model should ideally be assessed by external peer review by a group of individuals who were not involved in its development. The members of this group should not have a vested interest in the problem being addressed, or any knowledge of the model's results. In practice, this can be difficult to achieve and in many cases this means that the descriptive validity of a model must be determined by the decision makers who are its intended target audience (Mandelblatt et al, 1996; Weinstein, et al, 2001; Gray et al, 2011; Petrou & Gray, 2011; Eddy et al, 2012; Roberts et al, 2012).

The internal validity of a model is its ability to produce results that are consistent with the evidence that was used to construct it. A two-step process is generally used to establish this form of validity. In the first step of the process the model's logical structure, probability and other data are carefully checked against the source material from which they were obtained. If there are any formulas or equations in the model, the results they produce for a given set of input values are checked by manual calculation. In the second step of the process the model is tested using extreme parameter values (i.e. the lowest and highest possible values) to determine whether the resultant changes in the size and direction of the model's outputs are within expectations. While this process can be used to establish that a model is free of computational and data entry errors, it cannot be used to evaluate its structure or whether or not the outcomes it produces are accurate. The external validity of a model is the extent to which it produces results that are consistent with actual event data or, if this is not available, the results of other models that address the same problem. A validation is said to be independent, if the results of a model are consistent with those of other models that were built using data obtained from sources other than those that were used in the model's construction. A validation is said to be dependent, if the results of a model are consistent with the results of other models that were constructed from the same data sources as were used in the model's construction. If the results produced by a model are inconsistent with those produced by other models then an analyst must be able to provide a reasonable explanation for this (Mandelblatt et al, 1996; Weinstein, et al, 2003; Gray et al, 2011; Petrou & Gray, 2011; Eddy et al, 2012).
The predictive validity of a model is its ability to accurately forecast future events. It provides a means through which a model can be independently validated, one that is relatively immune to the model being adjusted to better match observed results. While this form of validation can be extremely useful, models are usually developed to help inform timely decision making in situations where the available data is incomplete, and there is significant uncertainty. In most cases, predictive validity can only be demonstrated for models that have short term outcome measures in situations where it is possible to conduct a prospective clinical trial to collect the data required to do this, or where a research design for such as study has already been published but is yet to be conducted. Consequently, a model should not be criticized for its inability to accurately predict future events based on information that was not available at the time it was developed. A decision about whether the results of a model can be trusted, prior to it having demonstrated predictive validity, needs to be based on a consideration of the potential reduction in uncertainty that could be gained through the acquisition of additional data and the additional costs incurred and the benefits lost from a delay (Weinstein et al, 2003; Gray et al, 2011; Petrou & Gray, 2011; Eddy et al, 2012).
Conclusion

Decision analytic modelling is a systematic framework that is used to structure imperfect information about events and outcomes in situations where there is significant uncertainty, multiple competing objectives and a clear need to determine the option that returns the highest expected value. Its use in economic evaluation is underpinned by expected utility theory. It is a process through which the expected value of at least two alternatives is derived from the sum of costs and outcomes, and the probability of their occurrence.

The choice of the type of model, and its structure, is driven to a major extent by the characteristics of the disease and intervention under evaluation. Ideally the structure of a model should reflect the key features of the natural history of the disease of interest and be based on a series of events that reflect important biological and clinical events in the disease process and its treatment. Decision trees are the simplest and most widely used form of decision analytic modelling. They are of most use when the scenario they are used to represent is relatively uncomplicated and there is a limited number of chance events and little or no reoccurrence of health states. A disease that has a large number of chance events, or in which there is a frequent reoccurrence of health states is better represented through the use of Markov modelling. In this form of state transition modelling, diseases are broken down into a series of health-states, such as treatment, remission, progression and death, through which individuals transition between over fixed periods of time.

For a model to be useful to a decision maker they need to be confident that the results it produces are an accurate reflection of the impact that an intervention has on an outcome of interest. To promote confidence in a model, it needs to be transparent and to have undergone a process of extensive validation. A transparent model is one that is accompanied by a careful, detailed, non-quantitative description of its structure, parameters, equations and underlying assumptions, as well as a detailed technical description that could be used by an end user to build a working copy of the model, if they so desired. The validity of a model is how well it does what it was intended to do, and the accuracy its results. In most cases models cannot be directly validated, for if the data to do this was readily available then there would be little or need for the model. A model's validity is instead established through a process that is undertaken to demonstrate that it has descriptive, internal, and external and in some cases, predictive validity.
The next chapter of the thesis describes in detail the methods that were used a formal systematic literature review of the evidence for the effectiveness of the use of non-invasive domiciliary LTMV to support the respiratory function of individuals with Duchenne muscular dystrophy.
Chapter 5: Systematic literature review

Introduction

In the previous chapters of the thesis the reader was provided with an introduction to Duchenne muscular dystrophy, the underlying theory and methods of economic evaluation, and the use of modelling in economic evaluations. The role of this chapter in the thesis is to provide a summary of the evidence of the effectiveness of the use of non-invasive LTMV to support respiratory function in DMD, to determine the strength of that evidence, and to extract the best available estimate of effectiveness and other data for use in an economic evaluation.

As previously discussed, the evidence for the effectiveness of an intervention for use in an economic evaluation should ideally be obtained from a formal systematic literature review. The Centre for Reviews and Dissemination (CRD) guidelines for the systematic review of healthcare interventions were used as the methodological guide for a formal systematic literature review of the available evidence. In the review, a strict, pre-specified set of methodological principles were followed for the reporting of the literature searching techniques, the criteria used for the inclusion and exclusion of studies, the assessment of bias, and other issues.

The chapter begins with a brief overview of the history and development of LTMV, from the use of negative pressure ventilation during the poliomyelitis epidemics of the late 1930s, 40s and 50s, through to the current day and the use of non-invasive positive pressure ventilation delivered via a nasal mask. This is followed by an examination of the postulated mechanisms of action that underpin the use of LTMV in neuromuscular and chest wall diseases. The methods that were used in the formal systematic literature review are then described, along with its findings. The results of the review were reported in narrative form, as there was only one randomized controlled trial and the quantitative pooling of the findings of studies that use non-randomized research designs is not recommended. The chapter ends with discussion of the findings of the review, its limitations, and the impact that these limitations have on the economic evaluation.
The history and development of mechanical ventilation

Mechanical ventilation is the application of pressure, to the body, or to the airways, to support or replace inspiratory and expiratory muscle function. The earliest mechanical ventilators were non-invasive negative pressure devices that intermittently reduced the air pressure over the chest wall to less than atmospheric pressure. A reduction in pressure that: (i) expanded the chest wall; (ii) increased the volume of the thorax; (iii) decreased the pressure of gas in the lungs and (iv); triggered a flow of gas into the lungs to equalize the pressure differential (Metha & Hill, 2001; Simonds, 2001a; Bach, 2002b; Corrado & Gorini, 2002).

The Scottish physician John Dalziel is widely credited as the first person to build a negative pressure ventilator. In 1838 he constructed a device that enclosed a patient's body in a rigid cylinder with an air-tight seal through which only their head protruded, with a set of hand-operated bellows used to reduce the pressure within the cylinder. Dalziel's device was one of many negative pressure ventilators that were designed and used, in Europe and the US, prior to the development in 1928 of the Drinker Iron Lung, the first practical mechanical ventilator. The device, which was known as an iron-lung, was based around a one-ton iron cylinder with an electrically powered rotary blower. In 1931, John Emerson of Cambridge Massachusetts produced a simpler, less expensive, lighter and more convenient version of the iron lung, one that could be operated by hand during a power failure. The Emerson iron lung was widely used throughout the world to treat acute respiratory muscle paralysis during the poliomyelitis epidemics of late 1930s, 1940s and 1950s (Snider, 1989; Metha & Hill, 2001; Simonds, 2001a; Bach, 2002b).

While the use of negative pressure ventilation saved a large number of lives during the poliomyelitis epidemics, mortality remained high with 80% of cases of acute respiratory muscle paralysis ending in death. It was not until 1952, and an outbreak of poliomyelitis in Copenhagen, Denmark that any substantial reduction in mortality was seen. At that time there were only seven negative pressure ventilators in the city and at the peak of the polio epidemic there were as many as 70 patients a day that needed mechanical ventilation. This meant that many patients in respiratory failure were invasively ventilated via tracheotomies and volume cycled positive pressure ventilators that were normally used for the delivery of anaesthesia during surgery, or through the use of bagging circuits that were operated by medical students. There was an unexpected reduction in mortality of 50% observed in the patients who were invasively ventilated, which was found to be due to a reduction in the risk of aspiration pneumonia. As a consequence of this reduction in mortality there was a gradual move worldwide to the use of invasive positive pressure ventilation to treat acute and chronic respiratory muscle failure (Snider, 1989, Metha & Hill, 2001; Simonds, 2001a, Bach, 2002b).
Although invasive positive pressure ventilation has been shown to be an effective and reliable form of LTMV, it is expensive\textsuperscript{32} as its use requires: (i) relatively sophisticated equipment; (ii) a high level of technical expertise that in many countries can only be provided by appropriately qualified nursing staff and; (iii) the ongoing and frequent use of consumables such as suction catheters. There are also a range of complications that are a direct result of the need for a surgical procedure to create a tracheotomy, and the need for a tracheal tube. These complications take the form of: (i) acute injuries to the larynx, the esophagus and surrounding blood vessels during surgery; (ii) stomal infection; (iii) haemorrhage and; (iv) mediastinitis\textsuperscript{33}. The pressure exerted by the cuff of the tracheal tube on the walls of trachea can lead to the development of: (i) tracheomalacia\textsuperscript{34}; (ii) tracheoesophageal or tracheo-arterial fistulas; (iii) tracheal granulation tissue; (iv) tracheal stenosis; (v) tracheal obstruction and; (vi) chronic pain. The tracheal tube also provides a direct route to the lungs for bacteria and other foreign material which increases the risk of pneumonia (Bach, Ishikawa & Kim, 1997; Marieb, 1998; Metha & Hill, 2001).

The cost of invasive positive pressure ventilation and the complications resulting from the need for a tracheotomy ultimately limited its usefulness as a form of LTMV. Negative pressure ventilation continued to be the most commonly used form of LTMV until the early 1980s and the development of non-invasive positive pressure ventilation. Non-invasive positive pressure ventilation (positive pressure ventilation delivered via a nasal mask) was developed to address a number of problems associated with the long-term use of negative pressure ventilation, specifically: (i) musculoskeletal back and shoulder pain, as patients were required to always sleep on their backs; (ii) the need for assistance from up to two carers to enter the ventilator and apply the neck seal; (iii) limited portability, as even the lightest tank ventilators weighed more than 300 kg and; (iv) most importantly, their inability to synchronise ventilator activity with a patients breathing pattern, which increases the risk of obstructive apnoeas and oxygen de-saturations during sleep (Hill, 1993; Metha & Hill, 2001).

\textsuperscript{32} The cost of a maintaining an individual at home in the United Kingdom who requires continuous invasive ventilation (24 hours a day) has been estimated to be £ 150,000 per annum (€ 246,000 Euros, reference year not specified) (Simonds, 2001b).

\textsuperscript{33} Swelling & inflammation of the mediastinum, the area between the lungs (Fauci et al, 1998)

\textsuperscript{34} Weakness of the muscle wall of the trachea (Brooker, 2010)
Modern non-invasive pressure cycled, positive pressure ventilators are: (i) lightweight and portable; (ii) can be used without assistance; (iii) do not require that a patient sleep on their backs (iv); are able to synchronise ventilator activity to a patient’s breathing pattern and; (v) do not require a tracheotomy, sophisticated equipment, 24 hour nursing care or the frequent and ongoing use of consumables such as suction catheters. There are only a few minor problems associated with their use, all of which are mask related and easily treated, problems such as: (i) mouth and nasal dryness; (ii) eye irritation; (iii) conjunctivitis and; (iv) skin breakdown on the bridge of the nose (Leger et al, 1994; Metha & Hill, 2001; Simonds, 2001a; Goldberg, 2002; Schonhofer, 2002; American Thoracic Society, 2004; Toussaint, Chatwin & Soudon, 2007).

**Mechanism of action**

The mechanism of action for the short-term use of mechanical ventilation in neuromuscular diseases such as DMD is relatively straightforward; it increases transpulmonary pressures and inflates and ventilates the lungs. What is not clear is how the use of nocturnal LTMV normalises daytime arterial blood gases and respiratory function. There are three postulated mechanisms of action: (i) that it provides respiratory muscle rest, and through this improves muscle strength, by temporarily reducing or removing the load placed on weakened and fatigued muscles; (ii) that it resets the sensitivity of the chemoreceptor’s in the respiratory centre to arterial carbon dioxide and; (iii) that it restores pulmonary mechanics (Meyer & Hill, 1994; Leger & Hill, 2001; Nickol et al, 2005).

The first of these postulated mechanisms of action, respiratory muscle rest, is not generally well accepted. For while some studies have reported periods of sustained electromyographic silence, that is indicative of muscle unloading, others have shown no change. Furthermore, it is not clearly understood how the use of

---

35 The basic cyclic and regular rhythm of respiration is controlled by two clusters of neurons in a region of the brainstem known as the medulla oblongata. One of these sets of neurons, the dorsal respiratory group, cyclically fires off impulses via the phrenic and intercostal nerves at a rate of approximately 12-15 cycles per minute, which triggers a co-ordinated contraction of the diaphragm and external intercostal muscles. Each impulse triggers an inspiratory phase that lasts about 2 seconds, after which the muscles relax for an expiratory phase of 3 seconds. The most important influences on the rate of firing of these neurons, and consequently the rate and depth of respiration, are the levels of carbon dioxide and oxygen in arterial blood. The levels of these gases are constantly monitored by two sets of sensors known as chemoreceptors. The central chemoreceptors in the medulla oblongata monitor the levels of hydrogen ions in the cerebrospinal fluid (CSF) that are formed as carbon dioxide diffuses into the CSF from the surrounding blood vessels. The peripheral chemoreceptors in the aortic and carotid bodies of the heart respond primarily to changes in arterial oxygen levels. The central chemoreceptor respond rapidly to changes in the PH of the CSF, through the direct stimulation of the medullary respiratory neurons, to increase or decrease the rate and depth of respiration so that the partial pressure of arterial carbon dioxide levels are kept to within 37 to 43 mmHg. Under conditions of rest the level of arterial carbon dioxide is the primary stimulus for changes in respiratory rate and depth. The peripheral chemoreceptors only play a role in controlling respiratory rate and depth when arterial oxygen levels are abnormally low (Marieb, 1998).
mechanical ventilation inhibits active muscle contraction, or how respiratory muscle fatigue progresses to muscle failure. There are no generally accepted definitions of muscle fatigue, or methods for its measurement (Allen, 1996; Celli, 2001; Mehta & Hill, 2001; Simonds, 2001a; Elliott, 2002; Nickol et al, 2005; Toussaint, Chatwin & Soudon, 2007).

The second of the three postulated mechanisms of action is that the use of LTMV normalizes the hypercapnic ventilatory response of the chemoreceptors in the respiratory centers in the brain that control respiratory rate and depth. There is a reduction in the sensitivity of these receptors due to a maladaptive response to nocturnal hypercapnia, where the maintenance of ideal arterial blood gas levels is sacrificed to achieve and maintain REM sleep. Long-term mechanical ventilation reduces the renal retention of bicarbonate\textsuperscript{36}, and in doing so restores, the response of the chemoreceptors to hypercapnia. In the presence of normal bicarbonate levels, the chemoreceptor's of the respiratory center increase respiratory rate and depth as required to restore and maintain normal levels of arterial carbon dioxide. While the evidence in support of this mechanism is also inconclusive, it has the strongest evidence base of the three postulated mechanisms of action and is the most widely accepted (Elliott, 2002; Leger & Hill, 2001; Mehta & Hill, 2001; Nikol et al, 2005).

The last of three postulated mechanisms of action is that the use of LTMV restores pulmonary mechanics. In individuals with respiratory muscle weakness, pulmonary mechanics are abnormal because of the adoption of a shallow, rapid pattern of breathing that does not fully inflate the lungs. Areas of lung tissue that are not fully inflated collapse, and atelectasis\textsuperscript{37} develops which further reduces lung capacity and chest wall compliance which increases the work of breathing (the energy needed to breath). By fully inflating the lungs, LTMV restores pulmonary mechanics by clearing atelectatic lung tissue, improving chest wall compliance and decreasing the work of breathing. The evidence in support of the mechanism is also inconclusive. It is possible that one or all three of these postulated mechanisms is responsible for the improvement in daytime blood gases produced by LTMV, for the majority of the studies that have been conducted to date have only examined these mechanisms in isolation and consequently very little is known about how they relate to each other (Meyer & Hill, 1994; Metha & Hill, 2001; Elliott, 2002; Nickol et al, 2005; Toussaint, Chatwin & Soudon, 2007).

\textsuperscript{36} Bicarbonate is a weak alkaline which corrects acidosis by reducing the number of free hydrogen ions (Marieb, 1998).

\textsuperscript{37} Collapsed alveoli that are not involved in gas exchange (Brooker, 2010).
Systematic literature review

The Centre for Reviews and Dissemination (CRD) guidelines for the systematic review of healthcare interventions (Centre for Reviews and Dissemination, 2008) were used as the methodological guide for the conduct and reporting of the formal systematic review. As recommended in the CRD guidelines, prior to undertaking the review a search was conducted in November 2010 of the Cochrane library38 to determine if there were any existing or ongoing systematic reviews of the use of LTMV for the treatment of respiratory insufficiency in DMD. The following search terms were used: Duchenne; neuromuscular; and long-term mechanical ventilation. The search uncovered one existing review, that of Annane et al (2009).

The topic of the review of Annane et al (2009) was the use of nocturnal LTMV for the relief of hypoventilation related symptoms, and the prolongation of survival in individuals with neuromuscular or chest wall disorders. The studies that were included in the review were quasi-experimental or randomised controlled trials (RCTs) of the use of any type or mode of LTMV, for any form of neuromuscular or chest-wall disorder of any severity, with no exclusions made on the basis of either age or gender. The primary outcome measure in the review was the short or long-term reversal of hypoventilation related symptoms. The secondary outcome measures were: short-term or long-term reversal of daytime hypercapnia; improvements in lung function and sleep-disordered breathing; a reduction in unplanned hospital admissions and; one-year mortality. There were a total of 8 studies that met the inclusion criteria for the review. The review found that there was a relative reduction in the risk of death after the initiation of the LTMV of 0.62 (95% CI: 0.42-0.91). The authors concluded that evidence for the use of LTMV in neuromuscular and chest wall disorders was weak and that there was a need for large RCTs (Annane et al, 2009).

The broad range of neuromuscular and chest wall disorders in the review, the inclusion of multiple forms of mechanical ventilation (invasive, non-invasive etc.) and the lack of DMD specific outcome data meant that a new systematic review was required to produce an estimate of the effectiveness of the use of non-invasive LTMV in individuals with DMD. The objectives of the review were: (i) to systematically search for the best available evidence for the use of domiciliary non-invasive LTMV, delivered by a nasal mask to support respiratory function in individuals with DMD; (ii) to determine the strength of that evidence and; (iii) to extract the best available estimates of effectiveness, and other data, for use in an economic evaluation.

**Methods**

**Criteria for considering studies**

The following criteria were used to decide whether a study was eligible for inclusion in the review.

**Study design**

The study designs that were considered for inclusion in the review were: (i) randomised controlled trials (RCTs) and; (ii) studies, with a comparator, that used non-randomised research designs. Case studies\(^{39}\) and reports were excluded from the review due to the significant potential for bias\(^{40}\) in this type of research design. Ideally, a systematic review of the effectiveness of an intervention for use in an economic evaluation should be based on the findings of RCTs. The use of blinding and the random allocation to treatment and control groups in an RCT minimises the likelihood of bias and confounding\(^{41}\) from known and unknown variables, and a consequence of this a well conducted RCT (with an appropriate sample size) is able to produce strong evidence of a causal link between an intervention and an outcome (Mandelblatt et al 1996; Demicheli, Jefferson & Vale, 2002; Centre for Reviews and Dissemination, 2008).

\(^{39}\) Studies without comparison groups, in which observations are reported for a individual or series of individuals (Centre for Reviews and Dissemination, 2008).

\(^{40}\) Systematic deviation from the true effect of an intervention due to weaknesses in research design, or in the analysis or interpretation of results (Centre for Reviews and Dissemination, 2008).

\(^{41}\) Factors other than the intervention under study that can influence study outcomes (Last, 2001).
The relative rarity of DMD makes it difficult for researchers to meet the sample size requirements of RCTs, and consequently there are very few RCTs of any intervention for individuals with the disease. It was therefore necessary to include studies with non-randomised research designs in the systematic review. In a non-randomised research design, the intervention that an individual receives is not determined by random allocation but by treatment decisions made by doctors and/or patients. Studies that use this form of research design are prone to bias, and consequently their findings are generally less reliable than those of a well conducted RCT (Hanratty et al, 2007; Centre for Reviews and Dissemination, 2008; Davidson & Truby, 2009).

There are different types of non-randomised research designs, for example: (i) prospective cohort studies, in which two groups of individuals with a disease, or at risk of developing a disease, are identified and followed forward over time to assess the impact of an intervention; (ii) retrospective cohort studies, in which individuals are identified from their past medical records and followed forward over time, either through the use of their medical records or through primary data collection and; (iii) case-control studies, in which two equivalent groups of individuals from the same population, one with an outcome of interest (cases) and one without (controls) are compared to determine if there is any evidence of association between an previous exposure (i.e. to an industrial chemical) and an outcome of interest (i.e. an unusual form of bladder cancer). If it is known that very few RCTs have been conducted of a given intervention then these studies can be included in a systematic review if they use some form of comparison group (Mandelblatt et al, 1996; Elwood, 1998; Yabroff & Mandelblatt, 2003; Hanratty et al, 2007; Centre for Reviews and Dissemination, 2008; Davidson & Truby, 2009).

**Population**

To be considered in the review study participants were required to have a diagnosis of DMD, but there was no requirement that this be confirmed by DNA analysis. While this is a potential confounder, it was anticipated that there would be very few (if any) studies in which DNA analysis was used to confirm the diagnosis of the disease. No restrictions were placed on the severity of the disease, or on the age of the study participants. Studies were excluded from the review if they had mixed populations. For example a range of neuromuscular diseases including DMD, and there was no reporting of DMD specific outcomes.
**Intervention**

To be considered for inclusion in the review, studies needed to examine the domiciliary use of non-invasive LTMV delivered via a nasal mask using a positive pressure ventilator. Consequently studies were excluded from the review if LTMV was delivered:

(i) During an inpatient admission in a hospital or other long term facility;
(ii) Using a form of mechanical ventilation other than non-invasive positive pressure ventilation, such as invasive positive pressure ventilation, delivered via a tracheotomy, or any form of negative pressure ventilation;
(iii) Through the use of multiple forms of ventilation (i.e. invasive ventilation, negative pressure ventilation & non-invasive positive pressure ventilation);
(iv) Via an interface other than a nasal mask, for example through the use of a mouthpiece seal and;
(v) As a component of a treatment regime that included the use of mechanically assisted coughing, as this is not widely used in Australia.

**Comparator**

To be considered for inclusion in the review studies needed to have a control or comparison group. This could take the form of a concurrent prospective control group who for example, received usual care that did not include the use of LTMV, or a retrospective comparison group who were not treated with LTMV.

**Outcomes**

The outcome measures that were examined in the review were: (i) the prolongation of survival and; (ii) quality of life. Studies were included in the review if they included one or both of these outcome measures. In a formal systematic literature review and in economic evaluations, the effectiveness of a therapeutic intervention is usually assessed in terms of clinically important patient relevant outcomes. These are outcomes that patients are aware of and wish to avoid, such as death, or events such as myocardial infarction or stroke that can severely impair physical functioning and quality of life (Drummond et al, 2005; Centre for Reviews and Dissemination, 2008; Taylor & Elston, 2009).

There is no clear evidence that a causal relationship exists between the normalisation of arterial blood gases and the prolongation of survival. Therefore, studies that used the normalisation of arterial blood gases as a surrogate measure for long term outcomes were not included in the review.
Language

To be considered for inclusion in the review studies needed to be published in English. This is the only language the author is able to read, and it was not possible to obtain full professional technical translations of papers due to a lack of financial resources. Ideally a systematic review should, where possible, include every study of an intervention that meets the inclusion criteria regardless of the language of publication. To not do so introduces the possibility of a language bias, as there may be studies with significant or non-significant results that have been published in languages other than English, that could change the conclusions drawn about the effectiveness of the intervention, if they had been included in the review (Centre for Reviews and Dissemination, 2008).

Search methods for identification of studies

Title and topic searches of MEDLINE (ISI); CINAHL (EBSCO); and EMBASE (EBSCO) were conducted in December 2010 using the following terms: muscular dystrophy; Duchenne muscular dystrophy; neuromuscular disorders; neuromuscular diseases; restrictive thoracic; non-invasive ventilation; non-invasive ventilation; positive pressure ventilation; long term mechanical ventilation; chronic respiratory failure; respiratory insufficiency; and nocturnal ventilation. Specific details of the search strategy used in these databases can be found in Table 8. To minimise the risk of publication bias, a Grey literature search that used the same search terms was conducted in Open SIGLE a database of Grey literature (http://opensigle.inist.fr/, last accessed January 2011). Details of the search strategy that was used in the Grey literature search can be found in Table 9. There were no exclusions made in any of the searches for: (i) the type of publication/study; (ii) the year of publication or; (iii) the language of publication.

42 Disorders or disease that limit chest wall or lung expansion and reduce total lung capacity, for example: abnormalities of the spine or thoracic cage; fibrosis of lung tissue or neuromuscular diseases (Brooker, 2010)
Table 8: Search strategy used in Medline, Cinahl & Embase

<table>
<thead>
<tr>
<th>Search</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Muscular dystrophy</td>
</tr>
<tr>
<td>2</td>
<td>Muscular dystrophy Duchenne</td>
</tr>
<tr>
<td>3</td>
<td>Neuromuscular disorders</td>
</tr>
<tr>
<td>4</td>
<td>Neuromuscular diseases</td>
</tr>
<tr>
<td>5</td>
<td>Restrictive thoracic disorders</td>
</tr>
<tr>
<td>6</td>
<td>1 or 2 or 3 or 4 or 5</td>
</tr>
<tr>
<td>7</td>
<td>Non-invasive ventilation</td>
</tr>
<tr>
<td>8</td>
<td>Non invasive ventilation</td>
</tr>
<tr>
<td>9</td>
<td>Positive pressure ventilation</td>
</tr>
<tr>
<td>10</td>
<td>Long term mechanical ventilation</td>
</tr>
<tr>
<td>11</td>
<td>Mechanical ventilation</td>
</tr>
<tr>
<td>12</td>
<td>Chronic respiratory failure</td>
</tr>
<tr>
<td>13</td>
<td>Respiratory insufficiency</td>
</tr>
<tr>
<td>14</td>
<td>7 or 8 or 9 or 10 or 11 or 12 or 13</td>
</tr>
<tr>
<td>15</td>
<td>Nocturnal ventilation</td>
</tr>
<tr>
<td>16</td>
<td>15 &amp; 6</td>
</tr>
</tbody>
</table>
Table 9: Search strategy used in OpenSigle

<table>
<thead>
<tr>
<th>Search</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Muscular+dystrophy</td>
</tr>
<tr>
<td>2</td>
<td>Duchenne+muscular+dystrophy</td>
</tr>
<tr>
<td>3</td>
<td>Neuromuscular+disorders</td>
</tr>
<tr>
<td>4</td>
<td>Neuromuscular+diseases</td>
</tr>
<tr>
<td>5</td>
<td>Restrictive+thoracic</td>
</tr>
<tr>
<td>6</td>
<td>Non-invasive+ventilation</td>
</tr>
<tr>
<td>7</td>
<td>Non-invasive+ventilation</td>
</tr>
<tr>
<td>8</td>
<td>Positive+pressure+ventilation</td>
</tr>
<tr>
<td>9</td>
<td>Positive+pressure+ventilation AND Duchenne</td>
</tr>
<tr>
<td>10</td>
<td>Long+term+mechanical+ventilation</td>
</tr>
<tr>
<td>11</td>
<td>Mechanical+ventilation</td>
</tr>
<tr>
<td>12</td>
<td>Mechanical+ventilation AND Duchenne</td>
</tr>
<tr>
<td>13</td>
<td>Chronic+respiratory+failure</td>
</tr>
<tr>
<td>14</td>
<td>Chronic+respiratory+failure AND Duchenne</td>
</tr>
<tr>
<td>15</td>
<td>Respiratory+insufficiency</td>
</tr>
<tr>
<td>16</td>
<td>Respiratory+insufficiency AND Duchenne</td>
</tr>
<tr>
<td>17</td>
<td>Nocturnal+ventilation</td>
</tr>
</tbody>
</table>

The results of all searches were entered into an Endnote X4 bibliographic database (Thomson Reuters, Carlsbad, CA). Duplicate entries were located and removed using the program’s duplicate search function and manual searching by the author.
Data collection and analysis

Selection of studies

The titles and abstracts of the papers in the Endnote bibliographic database were examined by the author for possible relevance to the review. Studies that clearly did not meet the inclusion criteria for the review were marked in this database as 1st level exclusions, with the reason for the exclusion. Studies that appeared to be of possible relevance to the review were flagged for retrieval for closer examination, and full copies were obtained of these papers. These were examined by the author, and flagged as either having met all the inclusion criteria for the review, or as 2nd level exclusions with the reason for the exclusion.

Data extraction

The extraction of study data was undertaken by the author. No contact was made with any the authors of studies to attempt to obtain missing data. The following set of data was extracted from all included studies: (i) aim/objectives; (ii) research design; (iii) inclusion and exclusion criteria; (iv) participant age; (v) the number of participants in the intervention and control groups and; (vi) the number of hours a day that LTMV was used.
Assessment of methodological quality of included studies

In a formal systematic review the methodological quality of included studies is assessed to determine the extent to which the findings of the studies can be reasonably assumed to be due to the intervention, and not due to flaws in the research design, or in the conduct of the study. To not do so puts into question the reliability of the findings of the review and any conclusions that are drawn about it (Centre for Reviews and Dissemination, 2008).

The methodological quality of the included studies was accessed using the following criteria:

1. Confirmation of the diagnosis:
   a. The methods used in the study to confirm the diagnosis of DMD, and whether this included DNA analysis;

2. The exclusion of individuals with clinical histories that were inconsistent with a diagnosis of DMD:
   a. Whether or not the study excluded potential study participants who had clinical histories that were inconsistent with a diagnosis of DMD, such as: a loss of ambulation after 13 years of age; or survival to the mid-twenties or later without the use of some form of LTMV;

3. Selection bias:
   a. Systematic differences in the characteristics of the intervention and control groups at baseline that could influence their ability to respond to the intervention, and their long term prognosis (Last, 2001; Centre for Reviews and Dissemination, 2008) and;

4. Performance bias:
   a. Systematic differences in the care provided to the intervention and control groups, other than the intervention of interest (Centre for Reviews and Dissemination, 2008).
Data synthesis

The results of the review were reported in narrative form. It was not possible to undertake a meta-analysis as there was only one RCT, and the quantitative pooling of the results of studies that use non-randomised research designs is not recommended (Centre for Reviews and Dissemination 2008).

Results

The systematic literature search produced a total of 5433 unique references. After closer examination by the author of the titles and abstracts of these references, there were 5376 first level exclusions (98.95%). Full copies were obtained of 57 papers for further examination, of which there were 51 (89.47%) second level exclusions. This left a total of six studies that contained content of possible relevance to the review, and on closer examination it was found that 2 (33.33%) of these studies, those of Leger et al (1994) and Simonds et al (1998) did not have comparison groups and they were consequently excluded. There were therefore only 4 (0.07%) studies that met all of the inclusion criteria for the review. There were no studies that met the inclusion criteria that examined quality of life as an outcome measure.

On the next page there is a Quorum\textsuperscript{43} flow chart of the results of the systematic literature search (Figure 4). The details of the excluded studies that had the prolongation of survival as an outcome measure are listed in Table 10 with the reason\textsuperscript{44} for their exclusion. The details of the excluded studies that had quality of life as an outcome measure are listed in Table 11, with the reason for their exclusion.

\textsuperscript{43} As per the guidelines developed by the QUORUM group for the reporting and presentation of the findings of a systematic review (Centre for Reviews and Dissemination, 2008)

\textsuperscript{44} The reason given for the exclusion does not mean that the study would have otherwise been included in review. For example, a study that used mixed modes of ventilation, may also have lacked a comparator, or have simply been a descriptive study of the use of LTMV in a given country.
Figure 4: Quorum flow chart

Potentially relevant studies identified in the literature search and screened for retrieval (n=5433)

Studies excluded (n=5376)
- No relevant content (n=5268)
- Case study/series (n=11)
- Foreign language (n=26)
- Inpatient use (n=3)
- Invasive ventilation (n=5)
- Letter to the editor (n=1)
- Mixed modes of ventilation (n=10)
- Mixed population (n=2)
- Mouthpiece ventilation (n=3)
- Neuromuscular diseases other than DMD (n=2)
- Review (n=31)

Studies retrieved for more detailed evaluation (n=57)

Studies excluded (n=51)
- Conference abstract (n=1)
- Case study/series (n=3)
- Editorial (n=1)
- Inpatient use (n=2)
- Invasive ventilation (n=2)
- Letter to the editor (n=6)
- Mixed modes of ventilation (n=11)
- Mixed population (n=11)
- NMD other than DMD (n=11)
- Review (n=3)

Studies of possible relevance to the review (n=6)

Studies included in the systematic review (n=6)

Studies withdrawn by outcome (n=2)
- Prolongation of survival (n=2)

Studies with usable information (n=4)
- Prolongation of survival (n=4)
- Quality of life (n=0)
Table 10: Excluded studies - Long term mechanical ventilation (n=140)

<table>
<thead>
<tr>
<th>Reason for exclusion</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conference abstract (n=1)</td>
<td>Konagaya et al (2005)</td>
</tr>
<tr>
<td>Inpatient use (n=4)</td>
<td>Swinburne, Fedullo &amp; Shayne (1988); Fukunaga et al (1993); Wijkstra, Avendano &amp; Goldstein (2003); Tatara et al (2005);</td>
</tr>
<tr>
<td>Review (n=30)</td>
<td>Yernault (1984); Knoblauch &amp; Walther (1990); Mallory &amp; Stillwell (1991); Braghiroli &amp; Donner (1992); Edwards &amp; Howard (1993); Raphael et al (1993); Simonds (1993); Paditz (1994); Simonds (1994); Claman et al (1996); Nosek &amp; Holmes (1996); Barthenl (1997); Klier (1997); Simonds (1997); Bonekat (1998); Raphael et al (1998); Raphael, Chevret &amp; Annane (1999); Simonds (2000); Bach (2002a); Corrado &amp; Gorini (2002); Langmack &amp; Make (2002); Leger (2002); Shneeerson &amp; Simonds (2002); Wedzicha (2002); MacDuff &amp; Grant (2003); Simonds (2003); Fauroux &amp; Lofaso (2006a); Annane et al (2009); Toussaint, Chatwin &amp; Soudon (2007); Dreher &amp; Windisch (2010)</td>
</tr>
</tbody>
</table>
Table 11: Excluded studies - Quality of life in LTMV (n=27)

<table>
<thead>
<tr>
<th>Reason for exclusion</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invasive ventilation (n=1)</td>
<td>Dreyer, Steffensen &amp; Pedersen (2010b)</td>
</tr>
<tr>
<td>Mixed modes of ventilation (n=3)</td>
<td>Miller, Colbert &amp; Osberg (1990); Bach, Campagnolo &amp; Hoeman (1991); Markstrom et al (2002)</td>
</tr>
<tr>
<td>Review (n=4)</td>
<td>Ordal &amp; Sanders (1994); Shneerson (1997); Windisch et al (2002); Windisch &amp; Criere (2006)</td>
</tr>
</tbody>
</table>

There were a total of 4 studies that were included in the review. A summary of these studies can be found in Table 12, which lists: (i) the type of research design used in the study; (ii) the interventions and comparators; (iii) the number of study participants in the intervention and comparison groups; (iv) the mean age of the study participants and the age range and; (v) the average number of hours per night that LTMV was used by the individuals in the intervention group (if reported).
Table 12: Summary of the included studies (n=4)

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Intervention</th>
<th>n</th>
<th>Comparator</th>
<th>n</th>
<th>Mean age (Yrs, range)</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raphael et al (1994)</td>
<td>RCT&lt;sup&gt;(i)&lt;/sup&gt;</td>
<td>SC&lt;sup&gt;(i)&lt;/sup&gt; &amp; LTMV</td>
<td>35</td>
<td>SC&lt;sup&gt;(iii)&lt;/sup&gt;</td>
<td>35</td>
<td>15.5 (11-33)</td>
<td>&gt; 6 hrs</td>
</tr>
<tr>
<td>Vianello et al (1994)</td>
<td>NRRD&lt;sup&gt;(ii)&lt;/sup&gt;</td>
<td>SC&lt;sup&gt;(i)&lt;/sup&gt; &amp; LTMV</td>
<td>5</td>
<td>SC&lt;sup&gt;(iii)&lt;/sup&gt;</td>
<td>5</td>
<td>20.1 (12-27)</td>
<td>&gt; 7 hrs</td>
</tr>
<tr>
<td>Eagle et al (2002)</td>
<td>NRRD</td>
<td>SC&lt;sup&gt;(i)&lt;/sup&gt; &amp; LTMV</td>
<td>24</td>
<td>SC&lt;sup&gt;(iii)&lt;/sup&gt;</td>
<td>134</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Eagle et al (2007)</td>
<td>NRRD</td>
<td>SC&lt;sup&gt;(i)&lt;/sup&gt;, SFS &amp; LTMV</td>
<td>27</td>
<td>SC&lt;sup&gt;(iii)&lt;/sup&gt;</td>
<td>20</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SC&lt;sup&gt;(i)&lt;/sup&gt; &amp; LTMV</td>
<td>14</td>
<td>SC&lt;sup&gt;(iii)&lt;/sup&gt;</td>
<td>39</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SFS only</td>
<td></td>
<td></td>
<td></td>
<td>Not reported</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SC&lt;sup&gt;(iii)&lt;/sup&gt; &amp; SFS</td>
<td></td>
<td></td>
<td></td>
<td>Not reported</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>39 neither&lt;sup&gt;(v)&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td>Not reported</td>
<td></td>
</tr>
</tbody>
</table>

Legend: (i) Randomised, controlled trial; (ii) Non-randomised research design; (iii) Standard care (iv) Spinal fusion surgery; (v) Did not receive either SFS or LTMV.
Randomized controlled research designs

There was only one study that used a randomised controlled research design (RCT). Raphael et al (1994) conducted a multicentre RCT of the initiation of LTMV prior to the onset of diurnal hypercapnia. The aim of the study was to test whether early intervention would slow the decline of respiratory function, and prolong survival in DMD. Over a five year period a total of 70 study participants were randomised into control and experimental groups of equal size. While there were some baseline differences between the two groups, such as a higher proportion of left ventricular hypokinesia in the experimental group (37% versus 23%), the authors stated that there were no significant differences. There was no individual in the study, at the time of baseline measurement that had diurnal hypercapnia or had previously used any form of mechanical ventilation. Individuals in the experimental group were treated with LTMV for at least six hours per night. Both groups received conventional treatment, which included the use of antibiotics, physiotherapy and inpatient mechanical ventilation for the treatment of acute chest infections. The main study outcome was overall survival.

The number of hours per night that LTMV was used was monitored through self-reported usage, and home visits by a respiratory specialist. The study was conducted as an intention to treat analysis, with overall survival calculated through the use of Kaplan-Meier survival analysis. An interim analysis at six months (June 1st, 1991) for 62 individuals showed a significant decrease in survival in the experimental group as compared to the control group (6 versus 1, p=0.005) and recruitment was stopped. At the end of the study (January 1st, 1992) there were 8 (22.86%) deaths in the experimental group, and 2 (5.71%) in the control group (p=0.05) with the underlying cause of the death generally poorly recorded. The relative risk of death in the experimental group was found to be 10 times higher than that of the control group (p=0.04). There was no evidence that the early initiation of LTMV slowed the decline of respiratory function or improved survival (Raphael et al, 1994).

Non-randomized research designs

There were 3 studies that used non-randomised research designs. Vianello et al (1994) undertook a study of the clinical course of 10 individuals with diurnal hypercapnia, five of whom had refused treatment with LTMV due to a belief that it would further decrease their quality of life. To ensure that LTMV was used correctly, the individuals in the treatment group were visited at least once a week by a respiratory specialist. At the end of the 2 year study period there were no deaths in the group treated with LTMV, and four deaths in the group who refused the intervention. The mean time to death for the untreated group was 9.7 (SD 5.8) months.
Eagle et al (2002) conducted a review of the medical histories of 183 individuals who had been treated at the Newcastle Muscle Centre (United Kingdom) between 1967 and 2002. The aim of the study was to determine whether the use of domiciliary LTMV had improved survival in DMD. Long-term mechanical ventilation was initiated when vital capacity fell to less than 1250mls, with the average duration of treatment being 5 years. There were 27 deaths in the group of individuals who had been treated with LTMV, three (11.11%) of which were due to severe cardiomyopathy, and these were excluded from the analysis of survival. The mean age of survival was calculated using Kaplan-Meier survival analysis, for a reference date of February 28th 2002. The mean age of death by cause: (i) for severe cardiomyopathy was 16.9 years ($n=22$, 95% CI: 15.23-17.97); (ii) for respiratory failure, in individuals not treated with LTMV, it was 19.29 years ($n=134$, 95% CI: 18.61-19.97) and; (iii) for respiratory failure, in individuals treated with LTMV, it was 25.3 years ($n=24$, 95% CI: 23.11-26.58).

Eagle et al (2007) conducted a review of the medical histories of 100 individuals born between the years of 1970-1990 who had been treated at the Newcastle Muscle Centre (United Kingdom). The aim of the review was to examine the impact on survival of the combination of LTMV and spinal fusion surgery (SFS). The review included: (i) 20 individuals who had only been treated with SFS; (ii) 27 who had been treated with both LTMV and SFS; (iii) 14 who had been only treated with LTMV and; (iv) 39 who had not been treated with either intervention. The criterion used for the initiation of LTMV in the individuals who received the intervention, was symptomatic nocturnal hypoventilation or a vital capacity of less than 600 mls. The intervention was initiated at a mean age of 17.4 years in the group treated with SFS and LTMV, and at a mean age of 18.2 years in the LTMV only group. The outcomes of the analysis were reported as the median age of death, which was calculated using Kaplan-Meier survival analysis. In the group who were treated only with LTMV, the median age of death was 22.2 years, with 35.6% surviving to 24 years of age. The median age of death in the group who were treated with SFS and LTMV was 30 years, with 84% surviving to 24 years of age. In the group who did not receive either treatment, only 11% survived to 24 years of age.
Methodological quality of the included studies

As previously discussed, the assessment of the methodological quality of the included studies was based on: (i) the confirmation of the diagnosis of DMD in the study participants; (ii) the exclusion of individuals who had clinically histories that were inconsistent with a diagnosis of DMD; (iii) selection bias and; (iv) performance bias. The results of the assessment are summarised in Table 13 using the grading system recommended in the CRD (2008) guidelines, specifically: (i) Yes, the criteria was clearly met; (ii) No, the criteria was clearly not met and; (iii) unclear, it was not possible for the reviewer to determine whether or not the criteria had been met (Centre for Reviews and Dissemination, 2008).

Table 13: Methodology quality of the included studies (n=4)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Confirmation of diagnosis</td>
<td>Unclear</td>
<td>Yes</td>
<td>Unclear</td>
<td>Yes45</td>
</tr>
<tr>
<td>Exclusion due to clinical history other than DMD</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Unclear</td>
</tr>
<tr>
<td>Selection bias</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Unclear</td>
</tr>
<tr>
<td>Performance bias</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Unclear</td>
</tr>
</tbody>
</table>

Confirmation of the diagnosis

In the studies of Raphael et al (1994) and Eagle et al (2002) the authors did not report how the original diagnosis of DMD had been made in the study participants, or whether they had attempted to independently confirm the diagnosis. In the study of Vianello et al (1994) the authors stated that the original diagnosis of DMD had been made using standard clinical, enzymatic, electromyographic, and histochemical criteria but it is not clear whether this included DNA analysis. It is also not known whether they attempted to independently confirm the diagnosis. There was only one study, that of Eagle et al (2007), that reported that the diagnosis of DMD had been confirmed by DNA analysis in the majority of the study participants (93%). Therefore, other than for the study of Eagle et al (2007), it cannot be determined with any certainty just how many of the individuals in these studies actually had DMD, as opposed to other

45 But not in every case
types of congenital muscular dystrophy, such as severe forms of Becker muscular dystrophy and Limb girdle muscular dystrophy (Type 21).

**Exclusion criteria**

While there were three studies, those of Raphael et al (1994); Eagle et al (2002) and Eagle et al (2007) that excluded individuals who were ambulant after 13 years of age, there were no studies that excluded individuals who had survived to their mid to late twenties without the use of any form of LTMV. Unassisted survival to this age is unlikely in individuals with DMD given the reported natural history of the disease. Individuals who are able to do this, most likely, have either a severe form of Becker muscular dystrophy or Limb girdle muscular dystrophy (Type 21) that has been misclassified as DMD. It is possible that that the decline of respiratory function in the severe forms of these two diseases may progress at much lower rate than it does in DMD, and that their inclusion in the studies may have resulted in an overestimation of the effectiveness of the intervention.

**Selection bias**

In the two studies that were conducted prospectively, the experimental and control/comparator groups may not have been equivalent at baseline. In the study of Vianello et al (1994) it is difficult to determine whether the baseline respiratory function of the treated and untreated groups were equivalent, as 1 of the 5 individuals in the untreated group refused to undergo respiratory function testing. In the study of Raphael et al (1994) there was a higher proportion of left-ventricle hypokinesias (13/35, 37%) in the treatment group at baseline, as compared to the control group (8/35, 23%).

**Performance bias**

In 2 of the 4 studies there was a systematic difference in the treatment provided to the individuals in the experimental /intervention and control/comparator groups. In the studies of Vianello et al (1994) and Raphael et al (1994) the use of LTMV was monitored at home by respiratory specialists. It is quite possible that as consequence of these home visits complications such as chest infections would be detected and treated much earlier than in the control group, which in itself could impact on the individual's probability of survival. In the study of Raphael et al (1994) chest infections in the individuals in the control group were treated with invasive ventilation (during an admission to hospital) at a much higher rate than in the experimental group (26% versus 9%).
In the studies of Eagle et al (2002) and Eagle et al (2007) there was a retrospective comparison of the outcomes achieved by individuals who were treated with LTMV (and/or spinal fusion surgery) and a group of “natural history controls” who had died before the intervention was introduced into clinical practice. A comparison group that may have resulted in an overestimation of the effectiveness of the intervention, for the increase in life-expectancy observed after the introduction of LTMV in the 1990s, as acknowledged by the authors, may have been at least partially due to a general improvement over time in the care provided to individuals with DMD. For example, improvements in care such as the routine vaccination for influenza and pneumococcal disease, and the early use of antibiotics and aggressive chest physiotherapy to treat chest infections (Mandelblatt et al, 1996; Eagle et al, 2007; Wagner, Lechtzin & Judge, 2007).

**Intervention effects**

**Prolongation of survival**

In the majority of included studies (3/4, 75.00%) the use of LTMV was associated with the prolongation of survival. There were two studies, those of Raphael et al (1994) and Vianello et al (1994), with relatively short periods of follow-up of 1 to 2 years. There were relatively few deaths (8/35, 22.86% and 0/5 respectively) in these two studies in the individuals who had their respiratory function supported by LTMV, and consequently they are of little value in the estimation of the impact of the intervention on long-term survival. The results of the study conducted by Eagle et al (2007) lacked detail, with only the median age of death reported for the various sub-groups in the study (without an associated range or other measure of variance). There was only one study, that of Eagle et al (2002) that reported (in sufficient detail) the mean life-expectancy of a group treated with LTMV and a group who were not treated, that could be used to determine the impact of the intervention on the prolongation of survival. In this study the use of LTMV was associated with a mean prolongation of survival of 6.0146 years (range of 4.50 to 6.61). The quantitative pooling of the findings of studies that use non-randomized research designs is not recommended and so the measure of the effectiveness of LTMV for the prolongation of survival for use in the economic evaluation was obtained from only one study, that of Eagle et al (2002).

46 The difference in the mean age of death from respiratory failure of 19.29 years (n=134, 95% CI: 18.61-19.97) in the group who did not receive LTMV and 25.3 years (n=24, 95% CI: 23.11-26.58) in the group who did.
Quality of Life

There were no studies that met the inclusion criteria that had quality of life as an outcome measure.

The collection of other data

The percentage of deaths due to severe cardiomyopathy, and the percentage due to respiratory failure, and the age at which these occurred were required for the construction of the decision-analytic model (that was used in the economic evaluation). This information, listed in Table 14, was obtained from a single study that of Eagle et al (2002) as this was the only study that reported this information.

Table 14: Other data collected for the decision-analytic model

<table>
<thead>
<tr>
<th>Data</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of deaths due to severe cardiomyopathy</td>
<td>14.75 % (n=27/183, 95% CI: 9.95%-20.74%)</td>
</tr>
<tr>
<td>Proportion of deaths due to respiratory failure</td>
<td>85.25 % (n=156/183, 95% CI: 79.26%-90.05%)</td>
</tr>
<tr>
<td>Mean age of death from severe cardiomyopathy</td>
<td>16.9 years (n=22, 95% CI: 15.23-17.97)</td>
</tr>
</tbody>
</table>
Discussion

The objective of the systematic review was to: (i) obtain the best available evidence of the use of non-invasive LTMV for the treatment of respiratory insufficiency in DMD; (ii) to determine the strength of that evidence and (iii); to extract the best available estimates of effectiveness, and other data, for use in an economic evaluation.

The systematic literature search identified a total of 5433 papers of possible relevance, of which 5376 (98.95 %) were excluded from the review. Full copies of the remaining 57 studies (1.05 %) were retrieved for closer examination, of which only 4 (0.07 %) were included in the review. The methodological quality of these studies was generally poor. The diagnosis of DMD was not independently confirmed in the majority of the studies, with only one study clearly reporting the confirmation of the diagnosis through the use of DNA analysis. There were no studies that excluded study participants who had clinical histories that were inconsistent with a diagnosis of DMD, specifically unassisted survival to the mid to late 20s. There were also differences in the experimental/treatment and control/comparator groups at baseline and differences in how the two groups were treated. The evidence for the effectiveness of the intervention for use in the economic evaluation, the proportion of cardiac and respiratory deaths and the mean age of death (by cause) was ultimately obtained from a single study, that of Eagle et al (2002).

The lack of evidence from a series of well conducted RCTs, and the methodological shortcomings of the included studies is such that the evidence for the effectiveness for the intervention is weak, and consequently there is insufficient evidence to draw any clear conclusions about the effects of the intervention. This is in itself, is a useful finding for a systematic review, but not one that should be interpreted as demonstrating that the intervention is ineffective (Centre for Reviews and Dissemination, 2008; McCabe, Edlin & Round, 2010).

One of the limitations of the review is that a single researcher decided on the exclusion and inclusion of studies and performed all of the data extraction. In a systematic review, this is normally undertaken by at least two researchers, who independently decide whether studies meet the inclusion or exclusion criteria for the review, with any disagreements resolved through discussion and consensus, and if necessary assistance from a third party. The researchers then go on to independently extract the data for the review from the agreed list of included studies, a process that improves the reliability and reproducibility of the review (Centre for Reviews and Dissemination, 2008).
Another further limitation of the review is the potential for language bias, due to the exclusion of 26 studies that were published in languages other than English. The exclusion of studies simply on the basis of the language of publication has the potential to throw into doubt the results of a review, and any conclusions drawn about the evidence for the effectiveness of an intervention. In this case, it is possible that one or more of the studies may have reported that the use of LTMV had no impact at all on survival, as studies conducted in non-English speaking countries that report non-significant results are less likely to be published in English language journals, than are studies that report statistically significant results (Centre for Reviews and Dissemination, 2008).

In recent years there have been questions raised about whether the evidence based guidelines developed by government bodies for reimbursement and funding decisions can be directly applied to medical devices. For while these guidelines appear to be generic in nature they have ultimately been developed for use with pharmaceuticals. For example, there is an assumption that it will be possible to obtain evidence of effectiveness from a series of large, well conducted double blinded RCTs, evidence that may not be readily available for a device such a mechanical ventilator. There are also a range of factors that need to be taken into account in studies of medical devices that have the potential to impact on its efficacy and effectiveness, that do not need to be considered in studies of pharmaceuticals. Factors, such as the learning curve associated with the use of a device, the underlying disease process of the end user and how they use the device, machine upgrades and the modifications that are made to devices over time by their manufacturers, such as software upgrades. It is unlikely that a series of large, well conducted RCTs of use of LTMV in DMD will ever be conducted, due to: (i) the relative rarity of the disease which makes it difficult to meet the sample size requirements of these studies; (ii) the need for a period of follow-up of approximately 15 to 20 years and the cost associated with this and; (iii) the existing evidence from observational studies for the use of LTMV in neuromuscular and chest wall diseases in general is such that the use of a non-treatment control group, one in which LTMV is not used, may be seen as a breach of medical ethics (Turkington & Elliot, 2000; Drummond, Griffin & Tarricone, 2009; McCabe, Edlin & Round, 2010).

Economic evaluations are conducted to provide decision makers with information that they can use to make timely decisions about the use of scarce health care resources under conditions of uncertainty. It is not unusual for decisions to be made about the allocation of healthcare resources when the best available evidence is weak, incomplete or imperfect and surrounded by significant uncertainty. Furthermore, a decision to fund an intervention for a rare disease, on the basis of the findings of an economic evaluation, remains an efficient use of resources under much higher levels of uncertainty than would normally be acceptable. As, all other things being equal, the opportunity cost of an intervention for a rare disease that
is ultimately shown by further research to be ineffective is small in comparison to that of an intervention for a common disease. The intervention would only be available to a small number of individuals and therefore the opportunities for benefit that could have been gained from the alternative uses of these resources would also be small. This is true, even when the ratio of costs to benefits for an intervention is much higher than would be usually acceptable, as the total cost is kept low by the small number of individuals who would receive it. Consequently, while it may be difficult to demonstrate with certainty that an intervention for a rare disease is cost-effective, there is a greater tolerance among decision makers for uncertainty in the results of an economic evaluation for these conditions, than would normally be the case. Notwithstanding, an economic evaluation based on the evidence obtained from this systematic review therefore cannot be seen to be anything other than exploratory in nature (Halpern, McKenna & Hutton, 1998; Kuntz & Weinstein, 2001; Yabroff & Mandelblatt, 2003; Drummond et al, 2005; Drummond, Griffin & Tarricone, 2009; McCabe, Edlin, & Round, 2010).
Conclusion

Mechanical ventilation is the application of pressure, to the body, or to the airways to support or replace inspiratory and expiratory muscle function. The earliest mechanical ventilators were non-invasive negative pressure devices, cylinders with electrically powered rotary blowers that were widely used to treat respiratory muscle paralysis in the polio epidemics of the 1930s, 1940s and 1950s. There was a gradual move to invasive positive pressure ventilation to treat acute and chronic respiratory failure following the successful use of invasive positive pressure ventilation during an outbreak of poliomyelitis in 1952, in Copenhagen, Denmark. Ultimately the cost of invasive positive pressure ventilation restricted its use as a form of LTMV and negative pressure ventilation continued to be the most commonly used form of the intervention, until the early 1980s and the development of non-invasive positive pressure ventilation.

A formal systematic literature review was conducted of the evidence for the effectiveness of the use of non-invasive LTMV in DMD. Of a total of 5433 papers of possible relevance to the review, 57 (1.05 %) were retrieved for closer examination, of these only 4 (0.07 %) met all of the inclusion criteria for the review. There were no studies of the health related quality of life of individuals with DMD that met the inclusion criteria. The findings of the review were reported in narrative form, as there was only one randomized controlled trial. The methodological quality of the included studies was generally poor; with the evidence for the effectiveness of the intervention for use in the economic evaluation ultimately obtained from a single study with a non-randomized research design, that of Eagle et al (2002). Consequently, an economic evaluation based on the evidence obtained from this systematic review cannot be seen to be anything other than exploratory in nature.

In the next chapter of the thesis, the reader is provided with an overview of the history and development of services to provide domiciliary LTMV to individuals with neuromuscular and chest wall disorders, and details of the model of service delivery that was used as a basis for the cost estimate in the economic evaluation.
Chapter 6: The provision of domiciliary LTMV

Introduction

In the previous chapters of the thesis the reader was provided with an overview of DMD, the underlying theory and methods of economic evaluation, the use of decision analytic modelling in economic evaluations and the limited evidence for the effectiveness of the use of LTMV to support respiratory function in the disease.

As previously discussed, there is very little known about either the cost or Cost-effectiveness of domiciliary LTMV for individuals with DMD and for neuromuscular diseases and chest wall disorders in general. In the majority of the states and territories in Australia the provision of domiciliary LTMV is undertaken on an ad-hoc basis by the respiratory departments of the major tertiary referral hospitals in the capital cities. The only formal centralised LTMV service in Australia is the state based Victorian Respiratory Support Service (VRSS). To date, the estimates of the cost of service provision that have been produced by the VRSS are not disease specific. It was therefore necessary to develop a model of domiciliary LTMV service delivery for individuals with DMD that could be used to provide a cost estimate for use in the economic evaluation. The aim of this chapter is to provide the reader with a detailed description of this model, and how it was derived.

To meet this aim the chapter begins with an overview of the history and development of services for the provision of domiciliary LTMV, starting with the programs developed in France, the United Kingdom, the United States of America and Australia for the ventilator dependent survivors of the polio epidemics of the 1940s and 1950s. This is followed by: an examination of the models of service delivery that are currently used for the provision of LTMV; the infrastructure and staffing requirements of these services, from the medical staff who prescribe and monitor treatment, through to technical staff who maintain the equipment, and the outreach nurses who provide onsite-support; the equipment requirements for individuals with DMD; the published estimates of the costs of service provision; the model of service delivery used by the Victorian Respiratory Support Service and; the model of service delivery that was used in the economic evaluation.
The history and development of domiciliary LTMV

The first country to develop a service to provide domiciliary long-term mechanical ventilation (LTMV) was France. Pilot programs were run in the 1960s in centres in Lyon and Paris to provide the ventilator dependent survivors of the polio epidemics of the 1940s and 50s with an opportunity to live outside of hospital. The success of these services led to the development, over the next 20 years of a further 25 centres. In 1991 the majority of these services were combined to form a national organisation, the Nationale pour le Traitement A Domicile de l’Insuffisance Respiratiorie Chronique (ANTADIR). In 1996 this organisation, with an annual budget of 92 million Euros, was responsible for the provision of LTMV and home oxygen to over 35,000 adults and children. France now has the largest and the best organized system in the world for the provision of domiciliary LTMV and the highest prevalence of its use (Donner et al, 2001; Leger & Hill, 2001; Make 2001; Stuart & Weinrich, 2004).

In the United Kingdom the development of domiciliary LTMV services dates back to 1965 and the Responaut program for polio survivors run by St Thomas’ Hospital, London. In the early 1980s the Royal Brompton Hospital, London and the Papworth Hospital, Cambridge began to develop services using negative pressure ventilators to provide LTMV for individuals with neuromuscular and chest wall diseases. In the mid to late 1980s there was a rapid expansion of the number of centres providing LTMV services following the development of non-invasive positive pressure ventilation (Goldberg, & Faure, 1984; Simonds, 2001b; Turner et al, 2003; Abbott, Carpenter & Bushby, 2009).

A detailed survey of the use of LTMV in 329 centres in 16 European countries was conducted by Lloyd-Owen et al (2005). In the survey the underlying reason for the need for the use of LTMV was divided into three categories: (i) lung: chronic obstructive pulmonary disease, cystic fibrosis, pulmonary fibrosis, pediatric diseases such as bronchopulmonary dysplasia and bronchiectasis; (ii) thoracic: obesity-hypoventilation syndrome, sequela of tuberculosis or lung resection such as thoracoplasty or lung resection and early onset kyphoscoliosis and; (iii) neuromuscular: motor neuron disease, muscular dystrophy, post-polio kyphoscoliosis, central hyperventilation syndrome, phrenic nerve paralysis and spinal cord damage. The mean prevalence of the use LTMV in the 16 countries (listed in Table 15) was estimated to be 6.6 per 100,000 on the 1st of July 2001 (Lloyd-Owen et al, 2005).
Table 15: Prevalence of LTMV in Europe

<table>
<thead>
<tr>
<th>Country</th>
<th>Number of centres</th>
<th>Prevalence (per 100,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>8</td>
<td>3.8</td>
</tr>
<tr>
<td>Belgium</td>
<td>23</td>
<td>5</td>
</tr>
<tr>
<td>Denmark</td>
<td>2</td>
<td>9.6</td>
</tr>
<tr>
<td>Finland</td>
<td>20</td>
<td>8.7</td>
</tr>
<tr>
<td>France</td>
<td>50</td>
<td>17</td>
</tr>
<tr>
<td>Germany</td>
<td>54</td>
<td>6.5</td>
</tr>
<tr>
<td>Greece</td>
<td>12</td>
<td>0.6</td>
</tr>
<tr>
<td>Ireland</td>
<td>15</td>
<td>3.4</td>
</tr>
<tr>
<td>Italy</td>
<td>70</td>
<td>3.9</td>
</tr>
<tr>
<td>Netherlands</td>
<td>4</td>
<td>5.6</td>
</tr>
<tr>
<td>Norway</td>
<td>38</td>
<td>7.8</td>
</tr>
<tr>
<td>Poland</td>
<td>8</td>
<td>0.1</td>
</tr>
<tr>
<td>Portugal</td>
<td>39</td>
<td>9.3</td>
</tr>
<tr>
<td>Spain</td>
<td>35</td>
<td>6.3</td>
</tr>
<tr>
<td>Sweden</td>
<td>65</td>
<td>10</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>40</td>
<td>4.1</td>
</tr>
</tbody>
</table>


The marked variation in the prevalence of LTMV in the 16 countries in the survey as listed in Table 15 (from 0.1 per 100,000 in Poland to 17 per 100,000 in France) is a result of differences in the rates of prescription for lung, neuromuscular diseases and chest wall disorders in these countries, and not the underlying prevalence of the conditions, for while there are mandated European Union guidelines for the prescription of long-term oxygen therapy there are no equivalent guidelines for LTMV. Furthermore, while there are a wide range of conditions for which LTMV can be used, there is very little rigorous evidence of its effectiveness and as a consequence of this, the conditions that it can be prescribed for in a given country are instead influenced by cultural, political and economic factors. For example, in Sweden respiratory specialists rarely ever prescribe LTMV for the treatment of smoking-related respiratory insufficiency, whereas in France LTMV is extensively used to treat this condition. There are also wide variations between countries in the resources available for the provision of LTMV due to differences in their national healthcare systems, for example, in Sweden and Denmark LTMV is provided free of charge to any individual that needs it, regardless of their level of healthcare insurance, or their personal financial...
resources. There are also marked variations between countries in the number of centres that can provide LTMV and the medical professionals who can prescribe it. For example, in Denmark LTMV can only be prescribed by the respiratory specialists of the two national LTMV centres, whereas in Switzerland it can be prescribed by any respiratory or intensive care specialist (Campbell & Pierce, 1998; Donner et al, 2001; Leger, 2001; Midgren et al, 2000; Laub, Berg & Midgren, 2004; Lloyd-Owen et al, 2005).

In the United States the development of services for the provision of LTMV was also driven by a need to care for the ventilator dependent survivors of polio. At the end of the 1940s there were approximately 400 ventilator dependent polio survivors being cared for in over 100 acute care hospitals across the country. In the early 1950s, to reduce the cost of this care, these individuals were moved into a small number of specialist regional LTMV centres. In 1953, to further reduce the cost of care and to improve quality of life, a domiciliary LTMV program was developed by the Rancho Los Amigos Hospital in California. The program was consequently implemented by all of the regional specialist centres, and by 1956 over 90% of the polio survivors were cared for at home. The development of effective vaccines against polio led to a decline in the number of cases of the disease and the need for specialist LTMV centres, and the majority of these centres were closed. Since that time domiciliary LTMV for neuromuscular diseases and chest wall disorders in the United States has become a rarity, and to date there is no national policy or system of funding for these services. Although in recent years there has been a revival of interest in the provision of domiciliary LTMV for children as a result of the availability of Medicaid funding for services for this population group (Downes & Parra, 2001; Simonds, 2001b; Bach, 2002b; Stuart & Weinrich, 2004).

In Australia, as in other countries, the development of services to provide domiciliary LTMV dates back to the polio epidemics of the late 1940s and 50s and programs run by Fairfield Hospital in Victoria and the Royal Prince Alfred Hospital in New South Wales. Since that time, while there have been editorials (Newton-John, 1989), journal articles (Gillis et al, 1989) and government reports (National Health and Medical Research Council, 1994) that have argued for the need for formal domiciliary LTMV services in Australia, little progress has been made and most services continue to be run on an ad-hoc basis (Gillis et al, 1989; Yates et al, 2004).

In the present day in Australia, domiciliary LTMV services for infants and children are run by the: Queen Elizabeth Hospital (South Australia); Westmead Hospital, Westmead (New South Wales); Children's Hospital, Camperdown (New South Wales); Royal Alexandra Hospital for Children (New South Wales); Mater's Children Hospital, Brisbane (Queensland); The Princess Margaret Hospital for Children, Perth (Western Australia) and; The Royal Children's Hospital, Melbourne (Victoria). There is no national
coordination of any of these services, or of the model of care they provide (Yates et al, 2004; Department of Human Services, 2005).

The service run by the Royal Children's Hospital, Melbourne (Victoria) ventilated its first patient in 1979, but it was not until 1985 that the state government of Victoria provided funding for the establishment of a formal program. The program provides all of the equipment required for domiciliary LTMV and pays for all ongoing consumables and attendant care. It also provides training and education for parents, and attendant carers, in the use of the equipment it supplies. The resources available to the service are limited and it is generally unable to provide LTMV to young adults with Duchenne muscular dystrophy (Health Solutions International, 1997; Jones et al, 2004; Department of Human Services, 2005; Tibballs et al, 2010).

In most states and territories in Australia domiciliary LTMV services for adults are run on an ad-hoc basis by the specialist respiratory medicine departments of the main tertiary referral hospitals, all of which are based in the capital cities. The only centralised formal LTMV service for adults is the state based Victorian Respiratory Support Service (VRSS). The VRSS was established in 1995 at the Austin and Repatriation Medical Centre (A&RMC) when a program run by the Fairfield Hospital for ventilator dependent polio survivors, and a program run by the A&RMC for neuromuscular and chest wall diseases, was combined into a single centralised service. The services provided by the VRSS include: the initiation and assessment of ventilation; training for ventilator users and their families; ventilator supply and maintenance; out-patient clinics for ventilator users; home visits by outreach nurses to monitor ventilator use and; emergency ventilator replacement and support. In 2004\(^47\) the VRSS provided LTMV services to 434 adults, of which 155 (35.71 %) were individuals with neuromuscular diseases and chest wall disorders (Health Solutions International, 1997; Campbell & Pierce, 1998; Jones et al, 2004; Department of Human Services, 2005).

\(^{47}\) The most recent year for which data was available
Models of service delivery

There are three broad models of service delivery that are used for the provision of domiciliary LTMV: (i) the ventilator and all necessary consumables, and follow-up care, are provided by a government funded public hospital or other government funded body (the majority of Scandinavian countries, the United Kingdom & Australia); (ii) the ventilator and all necessary consumables are provided to the patient by a private sector organisation based on a prescription from a hospital, with funding obtained from a variety of sources, with the prescribing hospital providing only follow-up medical care (Italy, Germany, Spain, Switzerland & the United States) and; (iii) the ventilator and all necessary consumables are provided by a private sector organisation, or a specialist community based not-for-profit organisation, with funding obtained from government (France & Canada) (Donner et al, 2001; Leger, 2001; Make, 2001).

There are common requirements for organisations that provide domiciliary LTMV, regardless of the model of service delivery they use, the diseases and conditions they manage, or the age of their target populations. These organisations need to have: (i) medical professionals who prescribe, initiate and/or monitor treatment; (ii) administrative staff who obtain funding, and organize the provision of ventilators for new users, and the on-going and regular replacement of masks, tubing and other consumables to existing users; (iii) technical staff who configure, and/or maintain the ventilators and other equipment and; (iv) outreach nurses who visit ventilator users at home, to undertake a general assessment of their health, and to determine whether the ventilator and mask are being used correctly. The specific roles and responsibilities of these individuals vary by country, for example: (i) in the United Kingdom, the Netherlands and Switzerland a respiratory specialist who prescribes LTMV is responsible for the on-going monitoring of its use; (ii) in France, after a prescription for LTMV is made by a respiratory specialist, its use is monitored by the doctors of the LTMV service providers and (iii); in the United Kingdom, the configuration, annual maintenance, repair and 24 hour emergency replacement of ventilators and other equipment is the responsibility of commercial organisations, who work under the direction of the hospitals who prescribe the intervention (Donner et al, 2001; Leger, 2001; Simonds, 2001b; Leger & Laier-Groeneveld, 2002; Nixon et al, 2008; Flunt, 2010a; Flunt, 2010b; Wark, Murray & Flunt, 2010).
There are two main types of ventilator that are used in the delivery of non-invasive domiciliary LTMV: volume cycled ventilators that deliver a pre-determined volume of air and; pressure cycled ventilators that deliver a flow of air until a pre-determined pressure is met. While volume cycled ventilators were originally developed to provide invasive positive pressure ventilation during surgery, they have been used for many decades to provide domiciliary LTMV. The relative complexity, weight and expense of these machines, and their inability to automatically compensate for the air leaks that inevitably occur with the use of nasal masks, has seen them being increasingly replaced by pressure cycled ventilators. Pressure cycled ventilators are, in comparison, relatively simple, lightweight and inexpensive devices that are able to automatically compensate for the air leaks that are associated with the use of nasal masks (Hill, 2001; Leger & Hill, 2001; Metha & Hill, 2001; Janssens et al, 2003; Lloyd-Owen et al, 2005).

In general when initiating LTMV the simplest available technology should be used, if an individual is able to breathe independently during the day, and is physically able to manipulate a nasal mask, there is little need for a ventilator to have alarms to alert the user about problems such as mask removal. An individual who is unable to manipulate a nasal mask, or who is unable to breathe independently at times during the day requires a ventilator with the following alarms: (i) mask removal; (ii) low pressure/disconnect and (ideally); (iii) power failure. While alarms such as mask removal were previously only seen in volume cycled ventilators, they are now standard features in modern pressure cycled ventilators. The majority of volume cycled ventilators have internal backup batteries in case of power failure, but this is not a standard feature in pressure cycled ventilators, although there are models that can be powered from a set of external batteries in emergencies or when travelling. Pressure cycled ventilators use much less power than volume cycled ventilators, and they are consequently able to run for long periods of time on relatively small batteries. A ventilator with alarms and battery backup, and a reserve machine, of the same brand and model with its own mask and tubing, is not strictly necessary unless an individual is ventilator dependent for more than 12 hours a day, or they are geographically isolated and a replacement ventilator cannot be provided in less than 4 hours (Simonds, 2001b; Leger & Laier-Groeneveld, 2002; Schonhofer & Sortor-Leger, 2002; Nixon et al, 2008; Flunt, 2010a; Flunt, 2010b).
The cost of service delivery

As previously discussed the cost of delivering domiciliary LTMV is very difficult to determine as there are few studies of either its cost or Cost-effectiveness. In many European countries there is no formal infrastructure for the provision of domiciliary LTMV and funding must be obtained on a case to case, ad-hoc basis. For example, in Italy there is no national policy or programme for the delivery of domiciliary LTMV and services are instead run by intensive care units, respiratory medicine departments and rehabilitation units. The capital cost of ventilators is met by the Italian national health care system, with technical assistance and support provided by either the equipment manufacturer (70% of cases) or by the biomedical engineering departments of the prescribing hospitals (30% of cases). The cost of assistance and support is met by the local hospital authority (90% of cases) or by the patients themselves (10% of cases), with the cost of consumables met by the prescribing hospital. In France, where domiciliary LTMV services are provided by private companies and non-profit organisations, the government reimburses service providers at a fixed rate that is dependent on the equipment requirements of an individual (as outlined in Table 16). Turner et al (2003) reported that the cost of initiating non-invasive LTMV in the United Kingdom was £ 3000\textsuperscript{48} for a pressure cycled ventilator, and £ 250 for two nasal masks and related consumables. The useable life of the ventilator, for an unspecified brand and model, was estimated to be 7 years, with two sets of nasal masks and related consumables required annually for each ventilator user. In an economic evaluation of the use of LTMV in motor neuron disease conducted by the National Institute for Health and Clinical Excellence (2010) the capital cost of a ventilator was reported to be £ 4000\textsuperscript{49} (for an unspecified type and model) with consumables costing £ 1713 per annum (Gasperini, Clini & Zaccaria, 1998; Donner et al, 2001; Leger, 2001; Leger & Laier-Groeneveld, 2002; Stuart & Weinrich, 2004).

\textsuperscript{48} UK pounds, reference year not known
\textsuperscript{49} UK pounds, reference year not known
Table 16: Reimbursement for LTMV in France

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Reimbursement50</th>
<th>Equipment / service provided</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-invasive ventilation</td>
<td>9 Euros a day</td>
<td>• A ventilator without alarms or battery</td>
</tr>
<tr>
<td>&lt; 12 hours a day</td>
<td>(€ 3285 a year)</td>
<td>• Three masks a year.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• A subsidy for electricity consumption</td>
</tr>
<tr>
<td>Non-invasive ventilation</td>
<td>10 Euros a day</td>
<td>• A ventilator with alarms and battery</td>
</tr>
<tr>
<td>&gt; 12 hours a day</td>
<td>(€ 3650 a year)</td>
<td>• Six masks a year.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• A second ventilator for users who are ventilated for more than 16 hours a day</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• A subsidy for electricity consumption</td>
</tr>
<tr>
<td>Invasive ventilation</td>
<td>15 Euros a day</td>
<td>• Ventilator with alarms and battery</td>
</tr>
<tr>
<td></td>
<td>(€ 5475 a year)</td>
<td>• Humidification, suction and safety equipment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Tracheotomy tubes, suction catheters</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• A second ventilator &amp; suction machine for individuals who are ventilated more than 16 hours a day</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• A subsidy for electricity consumption</td>
</tr>
</tbody>
</table>


The estimates of the cost of the provision of domiciliary LTMV in the United States come primarily from pilot programs, as until recently there has been no government funding available for this type of service. Dranove (1986) reported that it cost an average of $ 470 a day per child ($ 14,700 a month)51 for a group of 34 children who were cared for by the parents. This rose to $ 799 per day ($ 24,303 a month) for children who required professional nurses or other paid carers. Bach, Intintola, Alba and Holland (1992) estimated that domiciliary LTMV cost $ 235 a day ($ 7,050 a month) for group of 20 adults with severe neuromuscular disease. Sevick et al (1996) conducted a cross-sectional survey of 279 individuals with neuromuscular, spinal, congenital, and chronic obstructive pulmonary diseases, with the cost for these individuals ranged from $ 7814 to $ 40,678 a month. The wide variations in the cost estimates reported in these studies are due to a range of factors, such as: variations in the level of care provided; the severity of

50 Year not known
51 US$, for an unspecified reference year
the underlying disease; the presence of other disabilities; the need for professional nursing care and; the methods used to produce the cost estimates (Vick, 1996; Downes & Parra, 2001).

As previously discussed, the decentralised, mainly ad-hoc, nature of LTMV services in most states and territories in Australia make it difficult to estimate the cost of service delivery. In the majority of the states ventilators are issued by public hospitals or government equipment issuing centres using funding obtained from the federal government program of Aids for Disabled People (PADP). The criteria that have to be met to gain access to a government funded ventilator vary between the states and territories, in terms of the: (i) the conditions that can be treated; (ii) the equipment provided; (iii) the level of co-payment required; and (iv) who can be issued with equipment, with some states only providing equipment to low income earners and individuals who are on government pensions. In Victoria, the only state in Australia with a formal centralised LTMV service for adults, the supply of ventilators is funded through a special purpose grant from the Department of Human Services (DHS). The service delivery model used by the VRSS for the initiation, provision and management of LTMV is outlined in Table 17. The rate of replacement of ventilators, masks and other consumables is not known, and as previously discussed the cost to the VRSS of providing LTMV to individuals with DMD is not known (Campbell & Pierce, 1998; Department of Human Services, 2005).
Table 17: VRSS LTMV service delivery model

<table>
<thead>
<tr>
<th>Event</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation of ventilation</td>
<td>• Outpatient specialist respiratory physician review</td>
</tr>
<tr>
<td></td>
<td>• Respiratory function testing, arterial blood gases</td>
</tr>
<tr>
<td></td>
<td>• Overnight admission for a sleep study, with blood gas monitoring</td>
</tr>
<tr>
<td>Implementation of</td>
<td>• Inpatient initiation (3 nights)</td>
</tr>
<tr>
<td>ventilation</td>
<td>• Assessment by medical staff, respiratory physiotherapist</td>
</tr>
<tr>
<td></td>
<td>• Trial of mask types</td>
</tr>
<tr>
<td></td>
<td>• Daytime trial on ventilation</td>
</tr>
<tr>
<td></td>
<td>• Overnight titration sleep study, with blood gas monitoring and adjustment of</td>
</tr>
<tr>
<td></td>
<td>ventilation settings and mask</td>
</tr>
<tr>
<td></td>
<td>• Training of patients (&amp;/or carers) in ventilator management</td>
</tr>
<tr>
<td>Follow-up</td>
<td>• Outpatient specialist respiratory physician review</td>
</tr>
<tr>
<td></td>
<td>• Respiratory function testing, arterial blood gas measurement</td>
</tr>
<tr>
<td></td>
<td>• Review of ventilation as indicated by clinical assessment</td>
</tr>
<tr>
<td>Outreach services</td>
<td>• Ongoing on-site nurse visits for patient and family education / support,</td>
</tr>
<tr>
<td></td>
<td>training of carers, review of clinical status, ventilation settings and</td>
</tr>
<tr>
<td></td>
<td>equipment</td>
</tr>
<tr>
<td></td>
<td>• The provision of consumables</td>
</tr>
<tr>
<td></td>
<td>• Annual ventilator servicing and maintenance</td>
</tr>
<tr>
<td></td>
<td>• The provision of replacement equipment during servicing</td>
</tr>
<tr>
<td></td>
<td>• Emergency replacement of ventilators</td>
</tr>
</tbody>
</table>

Source: Department of Human Services (2005).
Service delivery model for the economic evaluation

As previously discussed the only formal centralised LTMV service in Australia is the state based Victorian Respiratory Support Service (VRSS). To date, the estimates of the cost of service provision that have been produced by the VRSS are not disease specific. It was therefore necessary to develop a model of domiciliary LTMV service delivery for individuals with DMD that could be used to provide a cost estimate for the economic evaluation (the details of this model can be found in Table 18). The model was based on the service delivery model used by the VRSS, but with the following changes: (i) the initiation of LTMV is undertaken on an outpatient basis, with a single overnight admission for a titration sleep study and; (ii) the emergency replacement of ventilators is undertaken by the equipment provider, as is annual ventilator maintenance and; (iii) titration sleep studies are conducted once every six months after the initiation of LTMV, as recommended in current treatment guidelines for DMD.

In the VRSS model of service delivery, as in many countries in Europe, when an individual meets the criteria for the initiation of LTMV they are admitted to a hospital to have a medical and respiratory review, fitting of a nasal mask, a daytime trial of ventilation, training in the use of the ventilator and a titration sleep study to adjust ventilator settings. This approach has significant resource implications as it requires a non-acute admission into acute hospital bed for a use that may not be reimbursable. The need for a hospital admission is also inconvenient for patients and their families as it usually means they will need to take time off work. A hospital bed may also not be available when needed, which can delay the initiation of LTMV. In the model of service delivery that was used in the economic evaluation, the initiation of LTMV was undertaken on an outpatient basis, as the VRSS is the only LTMV service in Australia that is known to be funded for inpatient initiation. The outpatient initiation of LTMV is therefore believed to better represent how the intervention is initiated in most states and territories in Australia. Furthermore, two recent studies of the outpatient initiation of domiciliary LTMV conducted by Lujan et al (2007) and Chatwin et al (2008) found that the outpatient initiation of LTMV, for patients who were medically stable, was as effective as inpatient initiation (Leger & Laier-Groeneveld, 2002; Lujan et al, 2007; Chatwin et al, 2008).
Table 18: Model of LTMV service delivery that was used in the economic evaluation

<table>
<thead>
<tr>
<th>Event</th>
<th>Requirements</th>
</tr>
</thead>
</table>
| Implementation of ventilation | • Outpatient daytime trial of mask types & training of patients (&/or carers) in the use of the ventilator and mask  
                                • Overnight admission for a titration sleep study with blood gas monitoring, and adjustment of ventilator settings & mask. |
| Follow-up<sup>52</sup>        | 1. Six monthly overnight admission for a titration sleep study, with blood gas monitoring and adjustment of ventilator settings as required |
| Outreach services             | • Out-reach worker home visit shortly after initiation to provide training and support.  
                                • Six monthly out-reach worker home visits to provide: (i) patient and family education & support; (ii) on-going training of carers; (iii) review of clinical status, ventilation settings and equipment and; (iv) to supply replacement masks, headsets, tubing and ventilator dust filters  
                                • 24 hour telephone support delivered by prescribing hospital  
                                • Emergency replacement of ventilators by the equipment manufacturer or their representatives |
| Maintenance and support       | • Annual on-site ventilator maintenance by the ventilator manufacturer, or their representative  
                                • A replacement ventilator is provided once every five years  
                                • An emergency power supply is provided when vital capacity falls below 20% of predicted for age & sex |

<sup>52</sup> In addition to the standard six monthly multi-disciplinary team review for a non-ambulant individual with DMD
The details of specific model of ventilator, and the masks and other consumables provided to individuals with DMD in the model of service delivery, and the rate at which they are replaced are listed in Table 19. The ventilator that was used in the model of service delivery, the Resmed VPAP IV, is a commercially available bi-level positive pressure device that has mask and pressure alarms. It is a base level model that does not have an internal battery but it can be powered from an external 24 volt battery pack, such as the Resmed Power Station II. In the model of service delivery a Resmed Power Station II is provided to all ventilator users when they develop diurnal hypercapnia (VC falls to below 20% of predicted for height, age and sex). The Resmed Mirage FX is a standard commercially available nasal mask designed for use with non-invasive positive pressure ventilators, such as the VPAP IV. The estimate of the working life of the ventilator and the external battery pack was obtained from the National Institute for Health and Clinical Excellence guidelines for use of LTMV for individuals with motor neuron disease (National Institute for Health and Clinical Excellence, 2010) and the bi-annual replacement of masks and other consumables was obtained from the report of Turner et al (2003a).

**Table 19: Capital equipment & the supply and replacement of consumables**

<table>
<thead>
<tr>
<th>Item</th>
<th>Details</th>
<th>Supply &amp; replacement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ventilator</td>
<td>Resmed VPAP IV bi-level positive pressure ventilator</td>
<td>On initiation of LTMV; replacement every 5 years</td>
</tr>
<tr>
<td>Mask &amp; headset</td>
<td>Resmed Mirage FX, mask and headgear</td>
<td>On initiation of LTMV; six monthly replacement</td>
</tr>
<tr>
<td>Tubing</td>
<td>1.8 metre ventilator tubing</td>
<td>Six monthly replacement</td>
</tr>
<tr>
<td>Air filters</td>
<td>Resmed VPAP air filters</td>
<td>Six monthly replacement</td>
</tr>
<tr>
<td>Emergency power supply</td>
<td>Resmed Power Station II - external lithium rechargeable battery pack &amp; charger</td>
<td>When VC falls to below 20% of predicted; replacement every 2 years</td>
</tr>
</tbody>
</table>

Conclusion

The first domiciliary LTMV services were developed to provide the ventilator dependent survivors of the polio epidemics of the 1940s and 1950s with an opportunity to live outside of hospital. Since that time these services have evolved to provide domiciliary LTMV to individuals with a wide range of neuromuscular diseases and chest wall disorders. In 2001, in Europe alone, there were 329 centres for the provision of domiciliary LTMV, across sixteen countries, in which the mean prevalence of the use of LTMV was estimated to be 6.6 per 100,000 (of the total population). There are marked variations in the prevalence of LTMV in these countries, from 0.1 per 100,000 in Poland to 17 per 100,000 in France, that are a result of differences in the rates at which the intervention is used in lung, neuromuscular and chest wall diseases and disorders. For while there are mandated European Union guidelines for the prescription of long-term oxygen therapy, there are no equivalent guidelines for LTMV as there is very little rigorous evidence for the effectiveness of the intervention, and consequently its use, within these countries, is influenced by cultural, political and economic factors.

As previously discussed the cost of delivering domiciliary LTMV is very difficult to determine as there are few studies of either its cost or Cost-effectiveness. A lack of a disease specific cost estimate for the provision of LTMV in Australia for individuals with DMD made it necessary to develop a disease specific model of service delivery that could be used to provide a cost estimate for use in the economic evaluation. In this model, which was based on the VRSS service delivery model, LTMV was initiated on an outpatient basis with the frequency of follow-up based on the current treatment guidelines for individuals with DMD.

The next chapter of the thesis describes in detail the alternative treatment pathways that were examined in the economic evaluation, the decision analytic model that was used in the analysis, and the small scale study that was undertaken to attempt to measure the impact that the use of non-invasive LTMV has on health-related quality of life.
Chapter 7: The economic evaluation

Introduction

The previous chapters of the thesis provided the reader with details of the use of domiciliary non-invasive positive pressure ventilation in DMD, the methods that are used in the economic evaluation of healthcare interventions, the limited evidence for the effectiveness of the intervention and the staffing and equipment requirements of services that provide domiciliary LTMV.

The role of this chapter in the thesis is to describe in detail the methods that were used to combine, in an exploratory economic evaluation, the events in the clinical treatment pathway for individuals with DMD, the evidence of the effectiveness of the intervention, and the equipment and staffing requirements of a service for the provision of domiciliary LTMV. The chapter also contains a detailed description of the methods that were used in a small scale study that was undertaken to collect health-related quality of life data for use in the economic evaluation, along with the findings of the study and an exploration of the possible reasons for its failure to recruit the minimum number of study participants that were needed to be able to return a statistical significant result. The overall aim of the chapter is to provide the reader with a detailed understanding of methods that were used in the evaluation and their limitations.

To achieve this aim the chapter starts with the details of the framing of the economic evaluation, the use of a decision-analytic model, the use of Cost-effectiveness and Cost-utility analysis, its aim, perspective and time horizon. The structure of the decision analytic model is then described in detail. This is followed by the event pathways for the two interventions and the current practice comparator, the methods that were used in the identification, measurement and valuation of outcomes and the identification, measurement and valuation of costs. The chapter ends with details of the model inputs and parameters, the values that were used in the base-case-analysis, the ranges that were examined in uncertainty and sensitivity analysis, the data sources from which this information was obtained and the steps undertaken to validate the model.
Design of the economic evaluation

As previously discussed, while the use of LTMV to support respiratory function in DMD is recommended practice there are still many doctors who do not routinely offer the intervention to their patients. This is partly due to a belief that it is not a cost-effective treatment and that the quality of life of a ventilator user is poor. There is actually very little known about the Cost-effectiveness of the use of LTMV in neuromuscular diseases and chest wall disorders in general, and to date there are no economic evaluations of the use of the intervention for individuals with DMD. There is also little known about the quality of life of individuals with DMD who are non-invasively ventilated through the use of a positive pressure ventilator (as documented in Chapter 5).

As previously discussed in the first chapter of the thesis, in the last twenty-five years there has been a rapid increase in the use of domiciliary LTMV to support respiratory function in individuals with neuromuscular diseases and chest wall disorders. During this time governments have come under increasing pressure to contain or reduce the cost of their healthcare systems, and to attempt to do this they have come to demand evidence of Cost-effectiveness as a key criterion on which to base new and recurrent funding decisions. The aim of the thesis was to produce an estimate of the Cost-effectiveness and cost-utility of the use of domiciliary long-term non-invasive positive pressure ventilation to support respiratory function in non-ambulant individuals with DMD in Australia, that could be used as a basis for informed decision making by clinicians and government funding bodies.

To achieve this aim the costs and benefits of two alternative treatment pathways for the timing of the initiation of LTMV in non-ambulant individuals with DMD were compared to the timing of its initiation in current practice. The alternative pathways were: (i) the initiation of LTMV from the onset of SDB in REM, without hypercapnia (from 15.5 years when VC falls to 60% of predicted – Intervention 1) and; (ii) the initiation of LTMV from the onset of SDB in REM, with hypercapnia (from 17.5 years of age when VC falls to 40% of predicted – Intervention 2). As previously discussed, there are no clear and consistent guidelines for the timing of the initiation of LTMV in DMD which makes it difficult to determine when the intervention is initiated in current practice. For the purposes of the evaluation, in the current practice comparator, LMTV was initiated at 18.5 years (when VC falls to 30% of predicted). The alternative treatment pathways outlined above differ for the pathway in the current practice comparator only in relation to the timing of initiation of the intervention, and as a direct consequence of this, the length of time over which the intervention is used.
For the purposes of the evaluation the timing of the initiation of the intervention, in the two alternative treatment pathways, was based on the mean bi-annual decline of vital capacity (VC) reported in the study of Hahn et al (1997), and the onset and staging of sleep disordered breathing reported in the study of Mellies et al (2003). In the model of service delivery that was used in the evaluation, domiciliary LTMV is initiated on an outpatient basis and delivered non-invasively through a nasal mask and the use of a pressure limited positive pressure ventilator. The target population of the evaluation was a hypothetical cohort of three identical groups of 100 non-ambulant Australian individuals with DMD. The time horizon of the evaluation was the life expectancy of the target population from 15 years of age, which based on Eagle et al (2002), was at best a maximum of 11.58 years.

The two forms of economic evaluation that were used in the evaluation, Cost-effectiveness analysis and Cost-utility analysis, are the two forms of economic evaluation that are recommended for use in submissions to Australian government bodies for health-sector funding applications, such as for the listing of drugs on the Pharmaceutical Benefits Scheme. A health sector perspective was used in the evaluation, and not a societal perspective, as LTMV in Australia is funded out of the fixed health sector budgets of state and federal governments (Department of Health and Ageing, 2008).

A decision-analytic model was used in the economic evaluation to determine the treatment pathway that returns the highest expected value. The model was developed by the author without external funding and without the use of expert opinion, for a single use application. Costs were reported in 2012 AU dollars. The outcomes produced by the model were quality adjusted life years (QALYs), condition specific quality adjusted life expectancy and total costs.

A decision to undertake the analysis through the use of a decision-analytic model, as opposed to a Markov or other form of model, was due to the limited data that was available for the construction of a model, and to ensure that the model was no more complex than was strictly necessary to answer the research question. The disease process and the impact of the use of LTMV to support respiratory function could have been represented in a Markov process model with five health states: (1) severe cardiomyopathy, unventilated; (2) severe cardiomyopathy, ventilated; (3) unventilated, without severe cardiomyopathy; (4) ventilated, without severe cardiomyopathy and; (5) death (an absorbing health state). For each six monthly cycle of the model, the following pay-offs would be need to be accumulated for each of the health states: (i) quality adjusted utility weighted quality of life; (ii) condition specific quality adjusted life and; (iii) costs. The initial distribution of the members of the cohort would be in one of the two unventilated health states, based on the proportion of individuals with severe cardiomyopathy (and those without severe cardiomyopathy) as reported in Eagle et al (2002). For each cycle of the model, for each of the non-
absorbing health states, the following transition probabilities would be needed: (i) from an unventilated health state to a ventilated health state; (ii) from an unventilated health state to death and; (ii) from a ventilated health state to death. The only allowable transitions within the model would be from a less severe (unventilated) to more severe (ventilated) health state, or to death an absorbing health state from which no further transitions would be possible. There would also be no allowable transitions for individuals with severe cardiomyopathy to a health state without severe cardiomyopathy, or for individuals without severe cardiomyopathy to a health state with severe cardiomyopathy.

The limited data available for the construction of the model would mean that there would no change in the probabilities for transitions between the health states for the majority of the cycles of the model. For example, ventilated individuals without severe cardiomyopathy would remain in the same health state for approximately 20 cycles (a period of ten years) before they transitioned into the death health state. There would also be no change in condition specific or utility weighted quality of life for the majority of the cycles of the model, with the available data limited to mean pre and post ventilation quality of life. Furthermore, as is normally the case in situations where Markov models are used, there are: (i) a limited number of possible health states in the disease process and no reoccurring events, or significant uncertainty about the likelihood, or the timing, of the initiation of the intervention or of outcomes such as death, due to either severe cardiomyopathy or respiratory failure and; (ii) due to the inevitable progression of the disease and the lack of a cure or any form of treatment, there is no possible direction of travel through the model other than from a less severe to a more severe health state (Gray et al, 2001; Briggs, Claxton & Schulpfer, 2006; Gray, Clarke, Wolstenholme & Wordworth, 2011).
For the purposes of the economic evaluation it was assumed that:

- All of the individuals in the model were fully compliant with the requirements of treatment and monitoring and regularly attended all of their outpatient appointments, sleep studies etc.;
- Other than the differences in the timing of the initiation of LTMV, in the alternative treatment pathways and the comparator, that the treatment provided to each of the three cohorts would be the same;
- As the approach used in the evaluation was that of incremental analysis, the healthcare resource usage and benefits gained from treatments such as; drug therapy for cardiomyopathy, scoliosis surgery and the use of nutritional supplements could be excluded from the analysis as they could be reasonably assumed not to be impacted upon by changes in the timing of the intervention, and consequently would not be expected to vary between the interventions and the comparator and;
- All-cause mortality could reasonably be excluded from the analysis, as the two most common causes of death in 15-28 year old Australian males are accidental injuries and poisoning (Australian Institute of Health & Welfare, 2005), two events that were extremely unlikely in a cohort of severely disabled, wheelchair bound individuals who are highly dependent on assistance from others to meet even their basic daily care requirements.

**Aim of the evaluation**

The aim of the economic evaluation was:

“To examine, in an exploratory analysis conducted from a health sector perspective with a lifetime time horizon, the Cost-effectiveness and cost-utility of two alternative treatment pathways for the use of domiciliary non-invasive positive pressure LTMV in DMD and their impact on quality of life, as compared to current practice”.

**Model structure**

A relatively simple longitudinal, deterministic decision analytic cohort model was used in the evaluation to represent the natural history of the disease and the impact that LTMV has on life expectancy, health related quality of life and healthcare resource usage for three identical hypothetical cohorts of 100 non-ambulant individuals with DMD.

As recommended by Briggs, Claxton & Schulpher (2006) a biological process, in this case the decline of respiratory function after the loss of ambulation, was used to structure the model and to reflect what is
known about the natural history of the disease. As previously discussed, the decline of respiratory function in DMD in the model was based on the findings of the study of Hahn et al (1997). The mean bi-annual decline of VC reported in this study (as listed in Table 3) was divided by four to produce a mean six monthly decline of VC that could be used for the timing of events in the model, such as the initiation of LTMV. The model was developed, tested and evaluated in TreeAge Pro 2009, a decision analytic modelling program produced by TreeAge Software Inc. (Williamstown, MA, USA). Microsoft Excel 2007 (Microsoft, Redmond, WA, USA) was used to calculate the total costs for the healthcare resource usage for the terminal nodes of the event pathways.

The estimate of the effectiveness of the intervention for the prolongation of survival in the model was obtained from a formal systematic literature review (as described in Chapter 5). As previously discussed there were only 4 studies that met the inclusion criteria for the review. The methodological quality of these studies was generally poor and ultimately the estimate of the effectiveness of the intervention was obtained from a single study, that of Eagle et al (2002). The findings of this study were used in the model to determine: (i) the proportion of deaths that were due to severe cardiomyopathy; (ii) the proportion that were due to respiratory failure; (iii) the mean age of death from severe cardiomyopathy; (iv) the mean age of death from respiratory failure and; (v) the duration of ventilation. For example, in Intervention 1 (in the base-case analysis) where LTMV is initiated at 15.5 years of age, there were: (i) 0.1475% of the cohort, who have severe cardiomyopathy, who live a mean of 0.50 years pre-ventilation (from their entry into the model at 15.0 years of age) and a mean of 1.4 years post-ventilation prior to their death at 16.9 years of age (16.9-15.5=1.4); and; (ii ) 0.8525% of the cohort, who do not have severe cardiomyopathy, who live a mean of 0.5 years pre-ventilation and 9.8 years post-ventilation prior to their death at 25.30 years of age (25.30-15.50=9.8).

There were no studies that met the inclusion criteria for the systematic review that had HRQoL as an outcome measure, and a small scale study was conducted by the author to attempt to collect this information. The estimate of healthcare resource usage in the model was based on the event pathways for the interventions and the comparator, which in turn were based on the current practice guidelines for the management of non-ambulant individuals with DMD, and the model of service delivery for the outpatient initiation and management of domiciliary LTMV that was described in detail in Chapter 6. There was no use of expert judgement.

The decision analytic model (Figure 3) that was used in the evaluation has one decision node, three pathways, 5 chance nodes and 8 terminal nodes. The three pathways that branch off the decision node reflect the major events in the natural history of the disease, and differences in the timing of the initiation of
LTMV for the two interventions and the current practice comparator. The first chance node in each of the arms of the model is the proportion of the cohort in which cardiomyopathy is so severe that it will result in death (as a primary cause) prior to death from respiratory failure and the proportion in which cardiomyopathy, while present, is not severe enough to result in death (as a primary cause) prior to death from respiratory failure. The second chance node in the pathway for interventions 1 and 2 is for proportion of the cohort with severe cardiomyopathy who have their respiratory function supported through the use of LTMV prior to their death from severe cardiomyopathy. The terminal nodes for the each of the pathways correspond to the proportion of deaths (as a primary cause) of the members of the cohorts that were due to: (i) severe cardiomyopathy where LTMV was not used to support respiratory function (terminal nodes 1, 4 & 7); (ii) severe cardiomyopathy where LTMV was used to support respiratory function (terminal nodes 2 & 5) and; (ii) respiratory failure where LTMV was used to support respiratory function (terminal nodes 3, 6 & 8). There are three payoffs attached to each of the terminal nodes: (i) condition specific HRQoL; (ii) generic utility weighted HRQoL and; (iii) mean total cost.

The members of the three identical hypothetical cohorts of 100 individuals with DMD enter the model at 15.0 years of age. They enter the model at this time, as based on the decline of respiratory function reported in Hahn et al (1997) and the onset of SBD reported in Mellies et al (2003), their mean VC is 63% of predicted (for height, age and sex) and within the next six months they will develop SBD in REM (without hypercapnia). The individuals in three cohorts move through the model, from left to right, following the natural history of the disease, as the use of LTMV has no impact on the age of death of individuals with severe cardiomyopathy, they are the first to reach a terminal node and to die and leave the model. While cardiomyopathy is present in the remaining individuals, it is not severe enough to result in death (as a primary cause) and their respiratory function is supported through the use of LTMV until their death from respiratory failure (as a primary cause).

The differences in the timing of the initiation of LTMV in the three arms of the model results in changes in: (i) the proportion of individuals with severe cardiomyopathy who have their respiratory function supported with LTMV prior to their death; (ii) the duration of the use of LTMV and the healthcare resources required to do this and; (iv) health related quality of life. It is important to note, that as previously discussed; (i) the use of LTMV to support respiratory function in individuals with severe cardiomyopathy has no impact on the decline of cardiac function and consequently there is no change in the age of death of these individuals and; (ii) the initiation of LTMV prior to the onset of diurnal hypercapnia (in Intervention 1 and 2) does not act to provide any additional survival benefit, beyond that provided in the comparator, it simply prolongs the duration of ventilation with the potential to improve or, at the very least to maintain, health related quality of life through the treatment of sleep disordered breathing.
A graphic representation of the decision-analytic model can be found on the next page (Figure 5). In which the structure of the model and its components are laid out, with examples of the formula’s that were used to calculate the costs and outcomes produced by each of the alternatives.

56 The definitions of the variables that were used in the model are listed later in the chapter in Tables 28, 29 & 30
Figure 5: The decision analytic model

![Decision Analytic Model Diagram]

- Costs
  - CP_NoSCM_NoVP=149029.
  - CP_NoSCM_VIP=416666.
  - CP_SCM_NoVP=181133.
  - Inv1_NoSCM_NoVP=102700.
  - Inv1_NoSCM_VIP=303074.
  - Inv1_SCM_NoVP=102700.
  - Inv1_SCM_VIP=737299.
  - Inv2_NoSCM_NoVP=103860.
  - Inv2_NoSCM_VIP=324374.
  - Inv2_SCM_NoVP=183133.
  - Inv2_SCM_VIP=431100.
- Deaths by cause
  - SCM_deaths=0.1475
- HRQoL
  - AGDL_NoVP=0.1349
  - AGDL_VIP=0.1727
  - MRF28_NoVP=0.8860
  - MRF28_VIP=0.5105
- Life years
  - CP_NoSCM_YnoV=3.5
  - CP_NoSCM_YV=6.8
  - CP_SCM_YnoV=1.9
  - Inv1_NoSCM_YnoV=0.5
  - Inv1_NoSCM_YV=0.8
  - Inv1_SCM_YnoV=0.5
  - Inv1_SCM_YV=1.4
  - Inv2_NoSCM_YnoV=2.5
  - Inv2_NoSCM_YV=7.90
  - Inv2_SCM_YnoV=1.60
  - Inv2_SCM_YV=0.
- Proportion ventilated
  - Inv1_SCM_V=1.
  - Inv2_SCM_V=0.
- Disc_Rate=0.00
Event pathways

Intervention 1: SDB in REM without hypercapnia

In first of the two alternative treatment pathways domiciliary LTMV is initiated when mean VC falls to below 60% of predicted and there is sleep-disordered breathing in REM, without hypercapnia. This typically occurs at 15.5 years of age.

The pathway of events for Intervention 1 was as follows:

Cardiomyopathy, severe, not ventilated

1. At 15.0 years of age there is an outpatient multi-disciplinary team review (a neurologist, cardiologist, dietician, social worker/psychologist and respiratory specialist) with respiratory function testing, respiratory muscle strength testing, peak cough flow testing, arterial blood gas analysis and an overnight inpatient diagnostic sleep study;
2. Once every six months there is an outpatient multidisciplinary team review etc., (as above);
3. This continues until death (at home) as a result of severe cardiomyopathy (as a primary cause) (Terminal node 1).

Cardiomyopathy, severe, ventilated

1. At 15.0 years of age there is an outpatient multi-disciplinary team review (a neurologist, cardiologist, dietician, social worker/psychologist and respiratory specialist) with respiratory function testing, respiratory muscle strength testing, peak cough flow testing, arterial blood gas analysis and an overnight inpatient diagnostic sleep study;
2. Domiciliary long-term mechanical ventilation is initiated at 15.5 years of age, at which time there is an outpatient multi-disciplinary team review (etc., as above), a daytime trial of mask types & training of patients (&/or carers) in the use of the ventilator and mask, and an overnight inpatient admission for a titration sleep study. A ventilator is provided, along with a mask & nasal seal, headset and tubing, and there is a home visit from an outreach nurse to provide onsite training and support for the ventilator user, and their family;
3. Once every six months there is an outpatient multidisciplinary team review (etc., as above), an overnight inpatient titration sleep study and a home visit from an outreach nurse, during which the ventilator user is provided with a replacement mask, nasal seal, headset and tubing. The air filters on
the ventilator are changed by the outreach nurse, the ventilator settings are checked, and a there is a general assessment of the health of the ventilator user;

4. The ventilator is serviced annually in the ventilator user's home;

5. This continues until death (at home) as a result of severe cardiomyopathy (as a primary cause) (Terminal node 2).

Cardiomyopathy, not severe, ventilated

1. At 15.0 years of age there is an outpatient multi-disciplinary team review (a neurologist, cardiologist, dietician, social worker/psychologist and respiratory specialist) with respiratory function testing, respiratory muscle strength testing, peak cough flow testing, arterial blood gas analysis and an overnight inpatient diagnostic sleep study;

2. Domiciliary long-term mechanical ventilation is initiated at 15.5 years of age, at which time there is an outpatient multi-disciplinary team review (etc., as above), a daytime trial of mask types & training of patients (&/or carers) in the use of the ventilator and mask, and an overnight inpatient admission for a titration sleep study. A ventilator is provided, along with a mask & nasal seal, headset and tubing, and there is a home visit from an outreach nurse to provide onsite training and support for the ventilator user, and their family;

3. Once every six months there is an outpatient multidisciplinary team review (etc., as above), an overnight inpatient titration sleep study and a home visit from an outreach nurse, during which the ventilator user is provided with a replacement mask, nasal seal, headset and tubing. The air filters on the ventilator are changed by the outreach nurse, the ventilator settings are checked, and a there is a general assessment of the health of the ventilator user;

4. The ventilator is serviced annually in the ventilator user's home;

5. When vital capacity falls below 1000 mls (20% of predicted for age and sex) an external emergency battery pack is provided to the ventilator user. All other activities continue as previously described\(^{57}\);

6. Five years after the initiation of ventilation, a replacement ventilator is provided, with the first ventilator kept as a spare. All other activities, i.e. six monthly respiratory function testing etc., continue as previously described, until the death of the individual (at home) as a result of respiratory failure (as a primary cause) (Terminal node 3).

---

\(^{57}\) Respiratory function testing and respiratory muscle strength testing end at 23.0 years of age, as respiratory function is so poor that it can no longer be measured (Hahn et al, 1997).
Intervention 2: SDB in REM, with hypercapnia

In the second of the two alternative treatment pathways domiciliary LTMV is initiated when mean VC falls to below 40% of predicted, and there is sleep-disordered breathing in REM, with hypercapnia. This typically occurs at 17.5 years of age.

The pathway of events in Intervention 2 was as follows:

Cardiomyopathy, severe, not ventilated

1. At 15.0 years of age there is an outpatient multi-disciplinary team review (a neurologist, cardiologist, dietician, social worker/psychologist and respiratory specialist) with respiratory function testing, respiratory muscle strength testing, peak cough flow testing, arterial blood gas analysis and an overnight inpatient diagnostic sleep study;
2. Once every six months there is an outpatient multidisciplinary team review etc., (as above);
3. This continues until death (at home) as a result of severe cardiomyopathy (as a primary cause) (Terminal node 4).

Cardiomyopathy, severe, ventilated

1. At 15.0 years of age there is an outpatient multi-disciplinary team review (a neurologist, cardiologist, dietician, social worker/psychologist and respiratory specialist) with respiratory function testing, respiratory muscle strength testing, peak cough flow testing, arterial blood gas analysis and an overnight inpatient diagnostic sleep study;
2. Once every six months there is an outpatient multidisciplinary team review etc., (as above);
3. Domiciliary long-term mechanical ventilation is initiated at 17.5 years of age, at which time there is an outpatient multi-disciplinary team review (etc., as above), a daytime trial of mask types & training of patients (&/or carers) in the use of the ventilator and mask, and an overnight inpatient admission for a titration sleep study. A ventilator is provided, along with a mask & nasal seal, headset and tubing, and there is a home visit from an outreach nurse to provide onsite training and support for the ventilator user, and their family;
4. Once every six months there is an outpatient multidisciplinary team review (etc., as above), an overnight inpatient titration sleep study and a home visit from an outreach nurse, during which the ventilator user is provided with a replacement mask, nasal seal, headset and tubing. The air filters on
the ventilator are changed by the outreach nurse, the ventilator settings are checked, and a there is a general assessment of the health of the ventilator user;
5. The ventilator is serviced annually in the ventilator user's home;
6. This continues until death (at home) as a result of severe cardiomyopathy (as a primary cause) (Terminal node 5).

Cardiomyopathy, not severe, ventilated

1. At 15.0 years of age there is an outpatient multi-disciplinary team review (a neurologist, cardiologist, dietician, social worker/psychologist and respiratory specialist) with respiratory function testing, respiratory muscle strength testing, peak cough flow testing, arterial blood gas analysis and an overnight inpatient diagnostic sleep study;
2. Once every six months there is an outpatient multidisciplinary team review etc., (as above);
3. Domiciliary long-term mechanical ventilation is initiated at 17.5 years of age, at which time there is an outpatient multi-disciplinary team review (etc., as above), a daytime trial of mask types & training of patients (&/or carers) in the use of the ventilator and mask, and an overnight inpatient admission for a titration sleep study. A ventilator is provided, along with a mask & nasal seal, headset and tubing, and there is a home visit from an outreach nurse to provide onsite training and support for the ventilator user, and their family;
4. Once every six months there is an outpatient multidisciplinary team review (etc., as above), an overnight inpatient titration sleep study and a home visit from an outreach nurse, during which the ventilator user is provided with a replacement mask, nasal seal, headset and tubing. The air filters on the ventilator are changed by the outreach nurse, the ventilator settings are checked, and a there is a general assessment of the health of the ventilator user;
5. The ventilator is serviced annually in the ventilator user's home;
6. When vital capacity falls below 1000 mls (20% of predicted for age and sex) an external emergency battery pack is provided to the ventilator user. All other activities continue as previously described;
7. Five years after the initiation of ventilation, a replacement ventilator is provided, with the first ventilator kept as a spare. All other activities, i.e. six monthly respiratory function testing etc., continue as previously described, until death (at home) as a result of respiratory failure (as a primary cause) (Terminal node 6).

58 Respiratory function testing and respiratory muscle strength testing end at 23.0 years of age, as respiratory function is so poor that it can no longer be measured (Hahn et al, 1997).
Comparator: Current practice

In the comparator domiciliary LTMV is initiated when mean VC falls to 30% of predicted. This typically occurs at 18.5 years of age.

The pathway of events in the current practice comparator was as follows:

Cardiomyopathy, severe, not ventilated

1. At 15.0 years of age there is an outpatient multi-disciplinary team review (a neurologist, cardiologist, dietician, social worker/psychologist and respiratory specialist) with respiratory function testing, respiratory muscle strength testing, peak cough flow testing, arterial blood gas analysis and an overnight inpatient diagnostic sleep study;
2. Once every six months there is an outpatient multidisciplinary team review etc., (as above);
3. This continues until death (at home) as a result of severe cardiomyopathy (as a primary cause) (Terminal node 7).

Cardiomyopathy, not severe, ventilated

1. At 15.0 years of age there is an outpatient multi-disciplinary team review (a neurologist, cardiologist, dietician, social worker/psychologist and respiratory specialist) with respiratory function testing, respiratory muscle strength testing, peak cough flow testing, arterial blood gas analysis and an overnight inpatient diagnostic sleep study;
2. Once every six months there is an outpatient multidisciplinary team review etc., (as above);
3. Domiciliary long-term mechanical ventilation is initiated at 18.5 years of age, at which time there is an outpatient multi-disciplinary team review (etc., as above), a daytime trial of mask types & training of patients (&/or carers) in the use of the ventilator and mask, and an overnight inpatient admission for a titration sleep study. A ventilator is provided, along with a mask & nasal seal, headset and tubing, and there is a home visit from an outreach nurse to provide onsite training and support for the ventilator user, and their family;
4. Once every six months there is an outpatient multidisciplinary team review (etc., as above), an overnight inpatient titration sleep study and a home visit from an outreach nurse, during which the ventilator user is provided with a replacement mask, nasal seal, headset and tubing. The air filters on the ventilator are changed by the outreach nurse, the ventilator settings are checked, and a there is a general assessment of the health of the ventilator user;
5. The ventilator is serviced annually in the ventilator user’s home;
6. When vital capacity falls below 1000 mls (20% of predicted for age and sex) an external emergency battery pack is provided to the ventilator user. All other activities continue as previously described;
7. Five years after the initiation of ventilation, a replacement ventilator is provided, with the first ventilator kept as a spare. All other activities, i.e. six monthly respiratory function testing etc., continue as previously described, until death (at home) as a result of respiratory failure (as a primary cause) (Terminal node 8).

The identification of outcomes

The outcomes that were examined in the economic evaluation were condition specific quality adjusted life expectancy and quality adjusted life years (QALYs). These are significant patient relevant outcomes that are also clearly of interest to clinicians and decision-makers, and as previously discussed; the most important long-term goals for the use of LTMV in DMD are the prolongation of life expectancy and the maintenance and/or improvement of quality of life. In keeping with recommended practice for economic evaluations a condition specific quality of life instrument and a generic multi-attribute utility instrument were used in the evaluation to capture unexpected effects, to allow comparisons to be made to the findings of other studies, and to maximise the likelihood that even small changes in HRQoL would be able to be detected (Mandelblatt et al, 1996; Janssens, 2001; Jones, Carone & Bertolotti, 2001; Eagle et al, 2002; Fiorenza, Vitacca & Cini, 2003; Toussaint, Chatwin & Soudon, 2007; Rajmil, Perestelo-Perez & Herdman, 2010).

The HRQoL of individuals with respiratory disease is generally measured in clinical practice, and in research studies, using generic health profiles such as the Sickness Impact Profile59 (SIP), the Nottingham Health Profile60 (NHP), and the Short-Form SF-3661 (SF-36). As previously discussed, while the use of these instruments means that comparisons can be made across a range of disease and conditions, they may lack sufficient sensitivity to detect small changes in disease specific HRQoL. There are three condition specific instruments that are commonly used to measure HRQOL in individuals with chronic

---

59 136 items covering the following domains: sleep and rest; eating; work; housework; recreation; ambulation; body care and movement; social interaction; alertness and; emotional behaviour and communication. It takes approximately 30 minutes to complete (Janssens, 2001).
60 45 items that are weighted to obtain 6 domain scores (sleep, pain, energy, physical mobility, social isolation and emotion). It takes approximately 10 to 15 minutes to complete (Janssens, 2001).
61 36 items covering the following domains: physical functioning; physical role; emotional role; pain; vitality; social functioning; mental and general health. It takes approximately 10 minutes to complete (Janssens, 2001; Jordan-Marsh, 2002)
respiratory insufficiency: (i) the St George Respiratory Questionnaire (SGRQ)\(^{62}\); (ii) the Chronic Respiratory Questionnaire (CRQ)\(^{63}\) and; (iii) the Maugeri Foundation Respiratory Failure Questionnaire (MRF28) (Janssens, 2001; Jones, Carone & Bertolotti, 2001; Last, 2001; De Blasio, 2004).

The SGRQ and CRQ were developed to measure HRQoL in individuals with chronic airflow limitation in conditions such as asthma and chronic obstructive airway disease (i.e. chronic bronchitis and emphysema), and while these instruments have been used in individuals with neuromuscular and chest wall diseases it is not what they were developed for, and questions have been raised about their validity in this disease population. The MRF28 was developed by Carone, Bertolotti, Donner & Jones (1999) to measure HRQoL in individuals with chronic respiratory insufficiency (CRI), the result of restrictive or obstructive disorders, and prior to and after the initiation of LTMV and/or long-term oxygen therapy. The instrument contains 28 multiple choice items that address issues of everyday activities, respiratory health, emotional status and invalidity that are considered to be core concerns for individuals with chronic respiratory insufficiency regardless of their age or gender. The MRF28 is self-administered and takes approximately 7 to 10 minutes to complete, with the majority of the multiple choice items requiring only a true-or-false response (Janssens, 2001; Jones, Carone & Bertolotti, 2001; Carone & Donner, 2002; Windisch et al, 2003a; Janssens et al, 2004).

The MRF28 was used to measure HRQoL in the small scale study as it was the only instrument that had been specifically designed to measure HRQoL in individuals with chronic respiratory insufficiency. A copy of the MRF28 can be found in Appendix 2. Permission to use the instrument was obtained from the developers.

**The MRF28 is not a preference based instrument and therefore the scores it produces cannot be used to generate QALYs for use in a Cost-utility analysis.**

The validity of the MRF28 was established by correlating the scores produced by the instrument with: (i) previously validated instruments (i.e. the SIP and the SGRQ) that were administered to individuals with CRI who were being treated with LTMV and/or oxygen therapy and; (ii) physiological variables such as arterial oxygen levels, exercise tolerance, oxygen flow requirements at rest and with exercise and

\(^{62}\) Self-administered, with 76 items that address the following areas: the frequency and severity of respiratory symptoms; the limitation of activities due to breathlessness; the social and psychological impact of respiratory impairment. Takes approximately 15 minutes to complete (Janssens, 2001).

\(^{63}\) Interview administered with 20 items covering the following areas: fatigue; emotional function and mastery and; dyspnoea. The questions are individualized to relate to daily activities indentified as areas of concern for each patient and consequently it cannot be used to make comparisons between the HRQoL of individuals (Janssens, 2001) and so is not suitable for use in an economic evaluation.
breathlessness. The reliability of the MRF28 was established through statistical analysis of its internal consistency and the stability of its test-retest scores. The scores produced by the instrument are normally distributed and cover a much wider range than the scores produced by the SGRQ, which suggests that for individuals with CRI that it may better able to discriminate between different levels of disease (Janssens, 2001; Jones, Carone & Bertolotti, 2001; Carone & Donner, 2002; Janssens et al, 2004).

The Assessment of Quality of Life, Version 1 (AQoL1) was used to measure generic utility weighted quality of life in the small scale study. At the time the study was conducted the AQoL1 was the only utility weighted HRQoL instrument that had been constructed using preference values obtained from a sample of the Australian population. The AQoL1 is widely used in Australian cost-utility analyses, and is an acceptable instrument for use in submissions to government funding bodies such as the Australian Pharmaceutical Benefits Scheme (Drummond et al, 2005; Hawthorne & Osborne, 2005: Australian Government Department of Health and Ageing, 2008).

The AQoL1 was developed by the Health Economics Group at Monash University to be a reliable and psychometrically valid instrument for use in economic evaluations. The descriptive system of the instrument was constructed from a framework developed by the World Health Organization for the classification of disability and impairment, and from a review of existing HRQoL instruments. The AQoL1 contains 15 questions and 5 sub-scales that address: illness; independent living; social relationships; physical senses and; psychological well-being. There are four response levels for each question, the first which corresponds to a “good health” state and the fourth to a “worst health” state. It is easy to administer, with respondents asked to evaluate their health state during the past week and takes approximately 5 to 7 minutes to complete (Hawthorne, Richardson, Osborne & McNeil, 1997; Hawthorne, Richardson & Osborne, 1999; Hawthorne & Osborne, 2005; Hawthorne, Richardson & Day, 2009).

Utility scaling for the AQoL1 was undertaken, using the time trade-off technique, in a within-strata random sample of 437 Australians. The internal consistency of the AQoL1 is 0.81 (Cronbach’s alpha), with the internal consistency of five sub-scales ranging from 0.52 (psychological well-being) to 0.86 (illness). Content validity was established through a comparison to several other HRQoL instruments (EQ-5D, HUI-III and the 15D) and the use of structural equation modelling. Concurrent validity was established through a comparison of the scores generated by the AQoL1 with those of a group of instruments that measured different aspects of HRQoL (i.e. mood, functional status and general health). The population norms and minimally important difference for the AQoL1 was determined using data collected from the 1998 South

---

64 The theoretical basis of the measurement of unobserved constructs in psychology, that provides a framework for the development of HRQoL instruments and their evaluation (Richardson, McKie & Baniola, 2011)
Australian Health Population Omnibus Survey of 3,010 individuals, and four other longitudinal studies. The mean population score for the instrument is 0.83 (SD: 0.20) with a difference of 0.06 suggestive of a minimally important difference, and a difference of 0.13 or greater suggestive of an important difference (Hawthorne, Richardson, Osborne & McNeil, 1997; Hawthorne, Richardson & Osborne, 1999; Hawthorne & Osborne, 2005; Hawthorne, Richardson & Day, 2009).

A copy of the AQL1 can be found in Appendix 3. Permission to use the instrument was obtained from the Health Economics Group at Monash University.

The measurement of outcomes

The systematic literature review was unable to locate any studies of the HRQoL of individuals with DMD that were suitable for use in the economic evaluation, and it was therefore necessary to undertake a small scale study to attempt to collect this data. The process of obtaining ethics approval for the study was time consuming due to the vulnerable nature of the study population, as a consequence of their age, the presence of cognitive impairment, the severity of their disease, and their dependence on medical care. Medical care that was provided to them, either directly or indirectly, by two of the author's PhD supervisors at two of the largest tertiary referral hospitals in Victoria (The Royal Children's Hospital, Melbourne and St Vincent's Hospital, Melbourne). Therefore to gain ethics approval the study had to be conducted in such a way as to remove any possible perception of coercion. To do this, recruitment and data collection was undertaken by mail that was originated from, and returned to, the School of Population Health at the University of Melbourne. There was no direct recruitment through either the Royal Children's Hospital or St Vincent's Hospital Melbourne and access to the details of the study participants was restricted to the author and his primary PhD supervisor. Ethics approval for the study was obtained from Human Research Ethics at the Royal Children's Hospital (EHRC 24100A, 24100B) and from the University of Melbourne (50057).

After ethics approval for the study was obtained, the author was given access to a database developed by the Department of Neurology at the Royal Children's Hospital (RCH) that contained the names, date of birth, and mailing addresses for every individual with DMD who had either been born, or treated at the hospital, between the years of 1975 and 2000. The Royal Children's Hospital is the main centre in Victoria for the diagnosis and treatment of children and adolescents, and therefore the database could reasonably be expected to contain the details of a significant proportion of the total population of individuals in Victoria with DMD. When the database was accessed in 2001 it contained the details of 166 individuals of which:

---

65 Renewal of the initial approval
5 (3.01%) were less than 6 years of age, and consequently extremely unlikely to be non-ambulant; 2 (1.2%) were female; 68 (40.96%) were known to have already died and; 9 (5.42%), who based on their age could be reasonably assumed to have already died. This left a total of 82 potential study participants, which given the exclusions listed above, was close to the expected prevalence of the disease in the state of Victoria of 99 to 141.

A personalised introductory letter, participant information and consent form, and a pre-paid self-addressed return envelope were sent to the 82 potential study participants. The participant information and consent form contained details of the study, and how they would be required to complete two quality of life questionnaires, and a short questionnaire about their use of LTMV (Appendix 4), once every second month for a six month period. A consent form was also provided for the parents or guardians of individuals who were under the age of 18, or who were otherwise unable to provide a valid informed consent. To maximise recruitment into the study, one month after the initial mail-out, a second set of forms were sent to every individual who had not yet responded. A recruitment notice for the study was also published in a newsletter produced by the Muscular Dystrophy Association of Victoria, a government funded service provider for individuals with DMD.

The final results of the two mail outs and the recruitment notice were as follows; of the estimated 82 potential study participants there were: (i) 15 individuals (18.29%) for whom the mailing details were incorrect; (ii) 4 individuals (4.88%) who informed the author (by phone, email or return post) that they were still ambulant and consequently outside of the target population of the study; (iii) 13 individuals (15.85%) who notified the author that they did not want to participate in the study and; (iv) 12 individuals (14.63%) who agreed to participate in the study and provided written consent. One of these 12 individuals failed to return any of their questionnaires, and consequently there were a total of 11 individuals who participated in the small scale study.

The reason for the low recruitment rate is not known, but it may have been due to: (i) the inability of the author to provide study participants with financial and other incentives; (ii) the level of commitment required, which may have been too much to ask of a group of severely disabled individuals; (iii) the clear statement made in the participant information and consent form, that study participants were unlikely to

---

66 Once ambulation is lost, respiratory function begins to rapidly decline, and it is not until after ambulation is lost that sleep disordered breathing develops.
67 The two females with DMD were not asked to participate in the study, as the disease is rarely ever seen in females and there are no females with DMD in any of the studies of the use of LTMV
68 No attempt was made to recruit these individuals into the study to reduce the possibility of distressing their families.
directly benefit from their involvement in the study and; (iv) the individuals having already been enrolled in another study and so they were unable (or unwilling) to participate in an additional study. It is not an uncommon occurrence for studies of healthcare interventions to have difficulty recruiting study participants, and to fail to recruit the number of participants that are required for a statistically significant result. A researcher therefore needs to be pessimistic and to assume, in the planning stages of a study, that recruitment will be extremely difficult and that they will need to do everything possible to maximise recruitment. Unfortunately there is little in the way of rigorous research evidence to guide the development of an effective recruitment strategy, and researchers must instead learn from the strategies used by other researchers, seek expert advice, and ultimately put together their own best guess. Ideally, if ethics approval can be obtained, a recruitment strategy should be tested in a pilot study prior to its use in a full scale study, and the findings of the pilot study used to refine the approach. While the use of financial incentives can improve recruitment, this can be difficult to obtain ethics approval for, as it can be seen as a form of coercion that may influence the participation of individuals from low socio-economic groups. In general the only form of direct payment to study participants that it is possible to gain ethics approval for, is the reimbursement of travel and other expenses. An added difficulty in the recruitment of individuals with a disability, or chronic illness, is the physical or cognitive burden that their participation in the study will place on them, a burden that is in addition to that already imposed on them by their disability or illness. Finally, regardless of the target population of a healthcare intervention, individuals are unlikely to participate in a study if they believe that they are unlikely to directly benefit from their involvement (Glasser, 2008; Salanitro, Estrada & Allison, 2008; Torgerson & Torgerson, 2008; Gliklick & Leavy, 2011; Reynolds, 2011).

A possible further issue in relation to this study, is that the irreversible decline of respiratory function in DMD, the development of sleep disordered breathing, and the need for LTMV are emotionally difficult issues that many individuals with DMD do not wish to reminded about. It is not unknown for individuals with DMD, who have clinically evident daytime respiratory insufficiency, to refuse to acknowledge that they have any difficulty breathing at all, or that one day they will need to be mechanically ventilated, as these are seen as pre-terminal events. Furthermore, many parents find it difficult to discuss the later stages of the disease with their sons, even though they are well aware of the natural history of the disease. They may instead hope that their son’s disease will progress differently from that of others, a cure will be found, or that a new treatment will be developed that will mean that LTMV will not be needed (Hukins & Hillman, 2000; Parker et al, 2005; Erby, Rushton & Geller, 2006).

---

69 In June 2000 there were estimated to be 2,359,041 males in the state of Victoria, given the estimated prevalence of DMD of 42 to 60 per million of the total male population, there would be 99 to 141 individuals with DMD (Riley &
The results of the small scale study

The mean age of the 11 study participants in the small scale study was 18 years (SD: 5.16 range 11 to 27). There were 8 individuals\(^{70}\) (72.73%) who did not use any form of LTMV, either prior to their entry into the study or during the six month period of data collection. There were three (27.27%) individuals who were already using LTMV prior to their entry into the study, who continued to use the intervention during the study. On their entry into the study the three individuals had been using LTMV for a mean of 4.33 years (SD: 2.89). There was one individual who started using LTMV at 17 years of age after a series of chest infections. There were two individuals who started using LTMV at 21 years of age due to the presence of sleep disordered breathing. There was one individual who reported the use of manually assisted coughing and there no individuals who were using mechanically assisted coughing.

There was a steady decline over the six month period of data collection in the number of questionnaires that were returned by the study participants (Table 20). The author did not have ethics approval to contact the study participants to determine the reason for this decline. It is possible that if nothing had changed since the previous set of questionnaires had been completed, that the study participants may have felt that there was no reason for them to complete and/or return a new set of questionnaires.

Table 20: Number of questionnaires returned by study participants for each mail out (\(n=12\))

<table>
<thead>
<tr>
<th>Form</th>
<th>1st mail out</th>
<th>2nd mail out</th>
<th>3rd mail out</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRF28</td>
<td>8 (66.67 %)</td>
<td>6 (50.00 %)</td>
<td>4 (33.33 %)</td>
</tr>
<tr>
<td>AqoL1</td>
<td>9 (75.00 %)</td>
<td>6 (50.00 %)</td>
<td>4 (33.33 %)</td>
</tr>
</tbody>
</table>

The collection of study data through the use of postal response questionnaires is a method that is widely used in social and epidemiological research. It is not unusual for studies that use this form of data collection to have a low response rate from potential study participants. It is also not unusual for a study where data is collected at multiple points over time, to progressively lose study participants with a resultant reduction in the effective sample size of the study. If the decline in the loss of study participants is not

---

\(^{70}\) In the process of revising this thesis for resubmission, the author rechecked all of questionnaires that were used in the small scale study, and in the process found that one individual (Study ID 006) had incorrectly been listed as having used LTMV. This meant that in the revised thesis, there were 8 individuals who were not using LTMV, and 3 individuals who were, as opposed to 7 individuals and 4 individuals respectively in the thesis that was originally submitted for assessment. As a consequence of the extremely small number of participants in the small scale study the movement of this one individual between the two groups resulted in the use of LTMV being associated with an improvement in HRQOL (\(n=3\)), as opposed to a decrease in HRQoL (\(n=4\)) as had been reported in the earlier version of the thesis.
random, and the reason for this loss is unknown, then it may not be possible to exclude the possibility of selection bias as an alternative explanation of the findings of a study. A number of strategies have been developed to improve the response rate for postal questionnaires such as; the use of financial incentives; pre-paid self-addressed envelopes (as used in this study) and; personalised questionnaires and letters (as used in this study). A systematic review of the use of postal response questionnaires conducted by Edwards et al (2002) found that the strategies associated with the highest response rates were: the use of financial incentives; short questionnaires; contacting study participants before their questionnaires were mailed out; making contact with study participants after they had received their questionnaires and; providing a second copy of the questionnaires to non-responders (Edwards et al, 2002; Puffer et al, 2004; Torgerson & Torgerson, 2008; Gliklich & Leavy, 2011).

As previously discussed, financial incentives were not offered in the small scale study, and in all likelihood it would not have been possible to gain ethics approval to do this. Contact was not made with study participants to improve the response rate for the questionnaires, as ethics approval was not sought to do this to reduce the perception of coercion. In hindsight, a balance should have been sought between the need to reduce the perception of coercion, and the need to collect an unbiased and complete set of study data.

The sample size requirements for the MRF28 are based on effect size, as the minimally important difference for the instrument is yet to be determined (Janssens et al, 2004). On this basis the sample size needed to produce a statistically significant result (two tailed, between groups, alpha of 0.5 and a power of 0.80) were: 64 per group for a small effect of 0.5 (a total of 128); 26 per group for a moderate effect of 0.8 (a total of 52) and; 17 per group for a large effect of 1.0 (a total of 34) (Cohen, 1988). The sample size required for the AQoL1 to be able to detect a minimally important difference of 0.06 or greater (standard deviation of 0.25) was 272 per group (alpha of 0.5, power of 0.80)\(^1\) and approximately 58 per group (alpha of 0.5, power of 0.80)\(^2\) for an important difference of 0.13 or greater (standard deviation of 0.25) (Davidson et al, 1986; Hawthorne, Gruen & Kaye, 2009).

The small scale study clearly failed to meet the minimum sample size requirements for the MRF28 and AQoL1 and it therefore did not have sufficient power to be able to detect a statistically significant difference between the two groups.

\(^1\) Two equal size groups, unpaired t-test (Davidson et al, 1986).
\(^2\) Two equal size groups, unpaired t-test (Davidson et al, 1986).
It is not an uncommon problem for studies conducted in rare disease populations to have difficulty in recruiting enough study participants to be able to return a statistically significant result. In many cases, this is a result of the prevalence of the disease, which can be so low, that a study that is restricted to a single hospital, state or country may have little or no chance of meeting its sample size requirements. To be able to do this may instead require a collaborative study that is conducted across a range of institutions, states and/or countries. For this type of study to produce useful results there needs to be a uniform accepted standard of care that is provided to all study participants, regardless of the institution, state or country they are treated in. This is currently not the case for neuromuscular diseases such as DMD, as there wide variations in the care provided to individuals with neuromuscular diseases due to a lack of widely accepted standards of care, especially in relation to the use of LTMV, the timing of its initiation and its mode of delivery (Butcher, 2007; Torgerson & Toregerson, 2008; Griggs et al, 2009; Forrest et al, 2011; Simoens, 2011).

At this point, the questions arises of why the economic evaluation was completed given the relatively weak evidence for effectiveness of the intervention, the uncertainty surrounding the delivery of domiciliary LTMV in Australia, and the failure of the small scale HRQoL study to recruit a statistically significant sample. As previously discussed, it is only through research that individuals with rare diseases can hope to gain equitable access to effective healthcare interventions. Unfortunately, as in this study, the best available evidence for the effectiveness of an intervention in a rare disease population may be nothing more than a small number of uncontrolled studies. It is unlikely given the existing evidence from observational studies for the use of LTMV in neuromuscular and chest wall diseases in general; that a large scale randomized controlled trial of the use of LTMV in DMD will ever be conducted. If it was possible to gain ethics approval for such a study, one that included a no-treatment control group, it would a very difficult study to recruit participants into given that one of its primary outcome measure would need to be the impact of the intervention on life expectancy (Turkington & Elliot, 2000; Schieppati et al, 2008; Drummond, Griffin & Tarricone, 2009; McCabe, Edlin & Round, 2010; Panju & Bell, 2010; Gliklich & Leavy, 2011).

Economic evaluations are conducted to inform timely decision making about the use of scarce health care resources under conditions of uncertainty, and it is not usual for such decisions to be made when the best available evidence is weak, incomplete or imperfect, and surrounded by significant uncertainty. The evidence for the effectiveness of an intervention for a rare disease, such as DMD, for use in an economic evaluation will never be as comprehensive, or as strong as the evidence that underpins diseases, such as type 2 diabetes mellitus, that affect a significant proportion of the population. There are no definite rules for the minimum level of evidence for the effectiveness of an intervention that must be present before an economic evaluation can be conducted. It is not unusual for a decision-analytic model to include data
from a range of sources of varying quality and to not exclude the use of data solely on the basis that it was not possible to demonstrate a statistically significant difference between two groups. This is allowable as long as: (i) the uncertainty surrounding the data is clearly documented; (ii) it is made clear that the results of the analysis need to be considered in relation to the quality of available data and; (iii) the impact that the use of this data has on the findings of the analysis are examined in sensitivity analysis. While these models may lack in scientific rigor, they provide a basis through which the best available evidence can be structured in a systematic and logical manner to better inform the decision making process for interventions for individuals with rare diseases (Halpern, McKenna & Hutton, 1998; Kuntz & Weinstein, 2001; Weinstein et al, 2003; Cooper et al, 2007; Drummond et al, 2009; McCabe, Edlin & Round, 2010; Panju & Bell, 2010).

The valuation of outcomes

The two HRQoL questionnaires were valued using the guidelines and methods recommended by their developers. Statistical analysis was undertaken using STATA 7 (STATA Corporation, College Station, USA).

MRF28

The scores for the MRF28 were calculated as follows: (i) a true response to one of the 28 questions was scored as 1, and a not-true response as 0; (ii) the total number of response items was reduced by 1 for any item which was not completed; (iii) if more than 5 items were not completed the questionnaire was not scored. The responses were then summed to produce a percentage of the possible total score, with higher percentages corresponding to a poor quality of life, and lower percentages to a better quality of life (Carone, Bertolotti & Jones, 1999; Janssens et al, 2004).

A total of 18 questionnaires were returned by the study participants, of which there were 3 (16.67%) that could not be scored as they were missing more than 5 items. If an individual had returned more than one questionnaire during the study period, the scores of their questionnaires were averaged so that they contributed only a single overall result to the group mean. There were only 8 individuals of the group who were not treated with LTMV, whose forms were able to be scored. The mean MRF28 score for this group was 0.1340 (SD: 0.1850, range 0-0.5185, 95% CI: 0.0036-0.5787). The mean score for the 3 individuals who were treated with LTMV was 0.3835 (SD: 0.0148, range 0.3750-0.4005, 95% CI: 0.0084-0.9057). The answers provided by the individual study participants, and their individual and average scores are reported in Tables 21 & 22. The difference in the mean scores of the two groups of 1.5442 (calculated
using the formula for Cohen's $D^{73}$ was indicative of a moderate to strong effect (Ferguson, 2009). Unfortunately given the failure of the small scale study to recruit the minimum number of study participants required to demonstrate a statistically significant result, this difference cannot be interpreted as demonstrating that the use of LTMV was associated with an improvement in HRQoL.

While the number of study participants in the small scale study makes it difficult to draw any firm conclusions about the adequacy of the MRF28 in this population, it does appears that based on the amount of missing data that the instrument may not have been well accepted (Brazier et al, 1999). It is possible that this may have been due to narrow focus of the instrument, and its concern with the impact that respiratory insufficiency has on the ability of individuals to perform the activities of daily living.

The quality of life of a non-ambulant individual with DMD is not simply a product of isolated respiratory impairment, but of the severe generalised muscle weakness they suffer from, that regardless of their respiratory function, acts to severely limit their ability to perform these activities. Furthermore, as previously discussed, it is not uncommon for individuals with DMD to deny having respiratory problems even when these are clinically evident, as respiratory impairment is seen as a pre-terminal event, and consequently there may have been some study participants who were unwilling to complete the MRF28 or who found it difficult to complete.

$^{73} d = \frac{(\text{mean of treatment group} - \text{mean of comparison group})}{\text{pooled standard deviation}}$ (Thalheimer & Cook, 2002).
Table 21: Individual results for the MRF28 (1 of 2)

| Form | 1.1 | 1.2 | 1.3 | 2.1 | 2.2 | 2.3 | 2.4 | 2.5 | 2.6 | 2.7 | 2.8 | 3.1 | 3.2 | 3.3 | 3.4 | 4.1 | 4.2 | 4.3 | 4.4 | 4.5 | 5.1 | 5.2 | 5.3 | 5.4 | 5.5 | 5.6 | 5.7 | % | Mean |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 01/01 | NA | NA | NA | NA | NA | NA | NA | F | F | F | T | T | T | F | F | F | F | F | F | T | T | F | F | F | NA | 19.2 | NA |
| 01/02 | F | F | F | F | F | F | F | F | T | T | T | F | F | F | F | F | F | F | T | T | F | F | F | NA | 18.5 | 18.9 |
| 01/03 | F | F | F | F | F | F | F | F | T | T | T | F | F | F | F | F | F | F | T | T | T | T | F | F | F | NA | 11.5 | NA |
| 02/01 | NA | NA | NA | NA | NA | NA | NA | F | F | F | F | T | T | T | F | F | F | F | T | T | T | F | F | F | NA | 11.5 | 11.5 |
| 02/02 | F | F | F | F | F | F | F | F | T | T | T | F | F | F | F | F | F | F | T | T | T | F | F | F | NA | 0 | 0 |
| 02/03 | F | F | F | F | F | F | F | F | T | T | T | F | F | F | F | F | F | F | T | T | T | F | F | F | NA | 0 | 0 |
| 03/02 | F | F | F | F | F | F | F | F | T | T | T | F | F | F | F | F | F | F | F | F | F | F | F | F | F | NA | 0 | 0 |
| 07/01 | F | F | F | F | F | F | F | F | T | T | T | F | F | F | F | F | F | F | F | F | F | F | F | F | F | NA | 0 | 0 |
| 08/01 | NC | NC | NC | NC | NC | NC | NC | NC | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | NA | 0 | 0 |
| 09/01 | F | F | T | T | T | T | T | T | F | F | F | F | F | F | F | T | T | T | T | T | T | F | F | F | NA | 51.9 | 51.9 |
| 12/02 | T | F | F | F | F | F | F | T | T | T | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | F | NA | 11.5 | 11.5 |

Legend: NA – Not applicable; NC – Not completed

74 101, 102, 103 are the first, second and third sets of the MRF28 questionnaire that sent out over the six month period of data collection to the study participant with a study ID of 01. While all of the study participants were sent three questionnaires, they were not all returned, for example ID 12 only returned the second MRF28 that was sent out.
Table 22: Individual results for the MRF28 (2 of 2)

<table>
<thead>
<tr>
<th>Form</th>
<th>6.1: My ventilator interferes with my life a lot</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>7.1: How is your general health?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.2: How is your breathing ability?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Not ventilated (n=8)</td>
</tr>
<tr>
<td>01/01</td>
<td>Not applicable</td>
<td>Very good</td>
</tr>
<tr>
<td>01/02</td>
<td>Not applicable</td>
<td>Good</td>
</tr>
<tr>
<td>01/03</td>
<td>Not applicable</td>
<td>Not completed</td>
</tr>
<tr>
<td>02/01</td>
<td>Not applicable</td>
<td>Very good</td>
</tr>
<tr>
<td>02/02</td>
<td>Not applicable</td>
<td>Good</td>
</tr>
<tr>
<td>02/03</td>
<td>Not applicable</td>
<td>Not completed</td>
</tr>
<tr>
<td>03/02</td>
<td>Not applicable</td>
<td>Poor</td>
</tr>
<tr>
<td>06/02</td>
<td>Not applicable</td>
<td>Not completed</td>
</tr>
<tr>
<td>06/03</td>
<td>Not applicable</td>
<td>Very good</td>
</tr>
<tr>
<td>07/01</td>
<td>Not applicable</td>
<td>Very good</td>
</tr>
<tr>
<td>08/01</td>
<td>Not applicable</td>
<td>Satisfactory</td>
</tr>
<tr>
<td>09/01</td>
<td>Not applicable</td>
<td>Very good</td>
</tr>
<tr>
<td>12/02</td>
<td>Not applicable</td>
<td>Satisfactory</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ventilated (n=3)</td>
</tr>
<tr>
<td>05/01</td>
<td>False</td>
<td>Good</td>
</tr>
<tr>
<td>05/02</td>
<td>False</td>
<td>Satisfactory</td>
</tr>
<tr>
<td>05/03</td>
<td>False</td>
<td>Satisfactory</td>
</tr>
<tr>
<td>10/01</td>
<td>False</td>
<td>Very good</td>
</tr>
<tr>
<td>11/01</td>
<td>False</td>
<td>Very good</td>
</tr>
</tbody>
</table>
A total of 19 AQoL1 questionnaires were returned by the 11 study participants. These were scored using the instructions provided by the developers of the instrument, with the responses to each of the questions scored from 1 to 4. Where there were multiple responses to a given question, or written answers: (i) if two responses were selected for one question, the response for the better health state was used, e.g. if for Question 11, B and C were selected, then this response was coded as B; (ii) if the responses selected are non-consecutive, e.g. A & C, then this was coded as B, the mean value of the two responses and; (iii) if a written comment was provided, instead of one of the supplied responses, then this was used as a guide for the assignment of a value, e.g. if the respondent wrote for Question 11 that they “find it difficult to hear in a crowd where other people are talking” then this was coded as B, as if the respondent had indicated that they have “some difficulty hearing”. The coded responses for the study participants were then entered into a STATA database, and the health state index scores for the study participants were calculated using a STATA program developed for this purpose by the Centre for Health Economics, Monash University, Australia (Hawthorne, Richardson & Day, 2009).

There are approximately 1.6 million possible health-states that can be represented by the descriptive system of the AQoL1 and consequently it was not possible for its developers to undertake to directly measure the utility associated with every health-state. The STATA program instead uses a multiplicative approach to model the five dimensions of the instrument and to produce the overall utility score. Through the use of this approach a score 0 and 100 is derived from the five independent utility scores produced by the “all worst” and “all-best” health states, which in turn is used to convert the individuals coded responses to their corresponding utility weights. There are 64 possible health states for each of the five dimensions of the instrument. The disutility for each of these health-states is determined by a series of multiplicative equations, and a look-up table of disutility values. There are only four of the five dimensions that are used in the calculation of the overall health-state utility score (independent living, social relationships, physical senses and psychological well-being) due to the first dimension, illness being more a measure of the level of severity of the individuals pathology, than of their HRQoL. The disutility scores for the four dimensions are converted to the corresponding utility values using the following formula (utility = 1 – disutility), which are combined in another equation to produce an overall utility score, that ranges from 0.0 (a state equivalent to death) to 1.0 (a state of perfect health (Hawthorne & Osborne, 2005; Hawthorne, Richardson & Day, 2009).
As with the results for the MRF28, the scores of individuals who completed more than one AQoL1 over the study period were averaged so that each individual contributed only a single overall result to the group mean. The mean AQoL1 health index utility scores for the 8 study participants who were not being treated with LTMV was 0.1349 (SD: 0.0478, 95% CI: 0.0982-0.1716). The mean AQoL1 health index utility scores for the 3 study participants who were treated with LTMV was 0.1727 (SD: 0.0322, 95% CI: 0.0927-0.2527). Details of the results for each of the study participants are listed in Table 23.

The mean health-index utility scores of the two groups were substantially lower than the AQoL1 population norm of 0.83 (SD: 0.20). The mean difference in HRQoL between the two groups of 0.0378 was insufficient to suggest that there was even a minimally important difference (Hawthorne & Osborne, 2005). Due to the failure of the small scale study to recruit the minimum number of study participants required to demonstrate a statistically significant result, this finding cannot be interpreted as demonstrating that LTMV has no impact on HRQoL. As with the results for the MRF28, the number of study participants in the small scale study makes it difficult to draw any firm conclusions about the adequacy of the instrument for individuals with DMD, but based on the response rate and the lack of missing data (Brazier et al, 1998) it appears the AQoL1 may have been more acceptable to the study participants than was the MRF28. It is possible that this may have been due to the AQoL1 addressing a broad range of issues, than the MRF28, in relation to the disabling effects of illness and disease.
Table 23: Individual results for the AQoL1

<table>
<thead>
<tr>
<th>Form</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
<th>Q11</th>
<th>Q12</th>
<th>Q13</th>
<th>Q14</th>
<th>Q15</th>
<th>Score</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not ventilated (n=8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/01</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>0.0641</td>
<td></td>
</tr>
<tr>
<td>01/02</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>0.1506</td>
<td></td>
</tr>
<tr>
<td>01/03</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>0.1109</td>
<td>0.1085</td>
</tr>
<tr>
<td>02/01</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0.1979</td>
<td></td>
</tr>
<tr>
<td>02/02</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.1661</td>
<td></td>
</tr>
<tr>
<td>02/03</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0.1660</td>
<td>0.1766</td>
</tr>
<tr>
<td>03/02</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>0.0440</td>
<td>0.0440</td>
</tr>
<tr>
<td>06/02</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0.2332</td>
<td></td>
</tr>
<tr>
<td>06/03</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0.1554</td>
<td>0.1943</td>
</tr>
<tr>
<td>07/01</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.1384</td>
<td>0.1384</td>
</tr>
<tr>
<td>08/01</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0.1727</td>
<td>0.1727</td>
</tr>
<tr>
<td>09/01</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0.1185</td>
<td>0.1185</td>
</tr>
<tr>
<td>12/01</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>0.0748</td>
<td></td>
</tr>
<tr>
<td>12/02</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0.1773</td>
<td>0.1261</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ventilated (n=3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>05/01</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0.1469</td>
</tr>
<tr>
<td>05/02</td>
<td>2</td>
<td>4</td>
<td>NC</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>NC</td>
<td>2</td>
<td>0.1469</td>
</tr>
<tr>
<td>05/03</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0.1144</td>
<td>0.1360</td>
</tr>
<tr>
<td>10/01</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0.1856</td>
</tr>
<tr>
<td>11/02</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0.1965</td>
<td>0.1965</td>
</tr>
</tbody>
</table>

Legend: NC – Not completed
The identification and measurement of resource use

The identification and measurement of healthcare resource use was based on the event pathways for the two interventions and the comparator. A micro-costing approach was used for measurement of the resources consumed for the provision of LTMV (i.e. ventilators and ventilator related consumables), with a gross costing approach used for the measurement of medical services and clinical and diagnostic testing. The measurement of costs was broken down into the following categories: (i) medical services and clinical and diagnostic testing; (ii) sleep studies and; (iii) ventilator provision. In the last of these three categories, ventilator provision, costs were further broken down into: (i) items of capital equipment (ventilators and emergency battery packs and their maintenance and; (ii) related consumables. As previously discussed (in chapter 6) based on manufacturers recommendations the ventilators were serviced annually, and replaced every five years and the emergency battery packs, as per manufacturer recommendations, were replaced every two years. In the cost estimates, there is no reuse of the ventilators and/or the emergency battery packs of individuals who had died and no allowance was made for any increases in the staffing levels or infrastructure in the hospitals in where LTMV was initiated.

On the next page, in Table 24, there is a list of the items of resource use that were considered in the evaluation and the number of each of these items that were consumed by the two interventions and the comparator, prior to and after the initiation of ventilation, for individuals who did not have severe cardiomyopathy (terminal nodes 3, 6 & 8). The items in the table are grouped together under the following headings: (i) medical services; (ii) clinical and diagnostic testing; (iii) sleep studies; (iv) ventilator provision; (v) maintenance and support and; (vi) consumables.

The following items of resource use were but not measured or valued, due to a lack of data: (i) patient time - the opportunity cost of the time spent by a patient while receiving treatment (Johnson et al, 1999); (ii) patient and family travel and other out of pocket costs –the costs incurred by patients and their families in the process of travelling to hospital for treatment (i.e. through the use of a privately owned vehicle, or the hiring of a taxi, or the fares paid to travel on public transport) or for child care, or meals (Johnston et al, 1999) and; (iii) Informal care – the costs incurred by family members in the process of providing care to patients, which includes the unpaid time spent in providing patients with assistance to perform the activities of daily living, and to meet the requirements of their treatment regimens (i.e. stretching, positioning, chest physiotherapy) and any financial outlays that are required to be able to do this (Johnson et al, 1999).
Table 24: Resource usage for the interventions, and the comparator

<table>
<thead>
<tr>
<th>Item</th>
<th>Intervention 1</th>
<th>Intervention 2</th>
<th>Comparator</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medical services, clinical and diagnostic testing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurology reviews</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Respiratory reviews</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Respiratory function &amp; respiratory muscle testing</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Arterial blood gas analyses</td>
<td>21</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>Cardiology reviews</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Dietician reviews</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Social worker / psychology reviews</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td><strong>Sleep studies</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnostic sleep studies (&lt;18 years of age)</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Diagnostic sleep studies (&gt;18 years of age)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Ventilation titration sleep studies (&lt;18 years of age)</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Ventilation titration sleep studies (&gt;18 years of age)</td>
<td>15</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td><strong>Ventilator provision, maintenance &amp; support</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resmed VPAP IV bi-level ventilators</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Resmed power station II (emergency power supply)</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Episodes of ventilator maintenance (1st machine)</td>
<td>9</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Episodes of ventilator maintenance (2nd machine)</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Outreach nurse visits</td>
<td>20</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td><strong>Consumables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resmed Mirage FX nasal masks &amp; headsets</td>
<td>20</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>Ventilator tubing</td>
<td>18</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>Air filters (1st machine)</td>
<td>19</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>Air filters (2nd machine)</td>
<td>9</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>
The valuation of resource usage

The consumption of healthcare resource usage was valued using market prices which were taken to be representative of their opportunity cost. The reference year for costing was 2012, with all costs expressed in Australian dollars. The costs and outcomes for each of the arms of the model are calculated at the start of the six month period prior to death, with death in all cases assumed to occur at some point in the following six months. As recommended by Siegel, Weinstein & Torrance (1996) a discount rate of 3% (and 5% for purposes of comparability) was used for both costs and outcomes in the base-case analysis, with the impact of other rates examined in sensitivity analysis.

As recommended, where they were available, national tariffs were used for the valuation of resource usage in the economic evaluation. The Medicare benefits schedule (MBS) was used in the evaluation to value: (i) medical services – services provided outside of a hospital service by a qualified medical practitioner, or on the behalf of medical practitioner; (ii) diagnostic procedures – pathology tests, imaging and other investigational and diagnostic procedures other than those performed as a component of an inpatient admission to hospital; (iii) Hospital, non-admitted patient services – services provided to patients at a hospital, such as in an emergency department or outpatient clinic, that is not associated with a formal admission to hospital (Commonwealth Department of Health and Ageing, 2002; Drummond et al, 2005; Glick et al, 2007).

75 To ensure that there is consistency and comparability in submissions to government bodies in Australia there are a set of standard sources for the valuation of medical and health related services for use in economic evaluations. For example, the Pharmaceutical Benefits Scheme (PBS), the Australian national formulary list of government subsided drugs is recommended for the valuation of prescription drugs. The use of this, and other sources of national tariffs is generally recommended in economic evaluations as they: (i) represent the cost to government of resource use; (ii) they are developed using a common methodology – at least within a given country; (iii) they are free or relatively inexpensive to access; and (iv) they can be used to cost most, if not all, of the medical services, diagnostic tests and drugs used in the delivery of a healthcare intervention (Commonwealth Department of Health and Ageing, 2002; Drummond et al, 2005; Glick et al, 2007).
There were no admitted hospital patient services in the event pathways for the interventions or the comparators. Hospital admitted patient services are the services provided to a patient with admitted patient status, following a formal admission to hospital, which includes all of the diagnostic, medical, allied health and other services they are provided with during their time in hospital (Commonwealth Department of Ageing, 2002). The only activities in the event pathway that were provided by hospitals were for outpatient non-admitted services (i.e. multidisciplinary team reviews) and diagnostic procedures (i.e. respiratory function testing, arterial blood gases, and overnight sleep studies).

The full details of the data sources that were used for the valuation of healthcare resource usage in the evaluation are listed in Table 25. The unit costs of each item of resource usage is listed in Table 26, and where relevant the item number for the event or activity as listed on the Medicare benefits schedule (i.e. item number, 12213, a diagnostic sleep study for an individual older than 18 years of age, $577.05, AU 2012).
Table 25: Data sources for the valuation of resource usage

<table>
<thead>
<tr>
<th>Resource item</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical, and allied health services</td>
<td>Australia Government Department of Health and Ageing (2012) Medicare benefits schedule[^76]</td>
</tr>
<tr>
<td>Respiratory function testing</td>
<td>Australia Government Department of Health and Ageing (2012) Medicare benefits schedule</td>
</tr>
<tr>
<td>Arterial blood gas testing</td>
<td>Australia Government Department of Health and Ageing (2012) Medicare benefits schedule</td>
</tr>
<tr>
<td>Sleep studies</td>
<td>Australia Government Department of Health and Ageing (2012) Medicare benefits schedule</td>
</tr>
<tr>
<td>Ventilators</td>
<td>Online store of CPAP Direct[^77]</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.cpap.com.au">http://www.cpap.com.au</a></td>
</tr>
<tr>
<td>External battery supplies</td>
<td>Online store of CPAP Australia[^78]</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.cpapaustralia.com.au">http://www.cpapaustralia.com.au</a></td>
</tr>
<tr>
<td>Ventilator maintenance and repair</td>
<td>10% of purchase price (NICE, 2010[^79]).</td>
</tr>
<tr>
<td>Outreach nurse services</td>
<td>Researcher estimate</td>
</tr>
<tr>
<td>Nasal masks and headsets</td>
<td>Online store of CPAP Direct</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.cpap.com.au">http://www.cpap.com.au</a></td>
</tr>
<tr>
<td>Replacement ventilator tubing</td>
<td>Online store of CPAP Australia</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.cpapaustralia.com.au">http://www.cpapaustralia.com.au</a></td>
</tr>
<tr>
<td>Replacement ventilator air filters</td>
<td>Online store of CPAP Australia</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.cpapaustralia.com.au">http://www.cpapaustralia.com.au</a></td>
</tr>
</tbody>
</table>

[^76]: Last accessed 8/8/2012  
[^77]: Last accessed 8/8/2012  
[^78]: Last accessed 8/8/2012  
[^79]: National Institute for Clinical Excellence (2010)
Table 26: Resource usage, unit costs and data sources

<table>
<thead>
<tr>
<th>Item</th>
<th>Unit cost</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medical services, clinical and diagnostic testing</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurology review</td>
<td>$ 42.20</td>
<td>MBS item number 105</td>
</tr>
<tr>
<td>Respiratory review</td>
<td>$ 42.20</td>
<td>MBS item number 105</td>
</tr>
<tr>
<td>Respiratory function and respiratory muscle testing</td>
<td>$ 136.05</td>
<td>MBS number 11503</td>
</tr>
<tr>
<td>Arterial blood gas</td>
<td>$ 22.60</td>
<td>MBS item number 13839</td>
</tr>
<tr>
<td>Cardiology review</td>
<td>$ 42.20</td>
<td>MBS item number 105</td>
</tr>
<tr>
<td>Dietician review</td>
<td>$ 61.10</td>
<td>MBS item number 10954</td>
</tr>
<tr>
<td>Social worker / psychology review</td>
<td>$ 61.10</td>
<td>MBS item 10956</td>
</tr>
<tr>
<td><strong>Sleep studies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnostic sleep study (&lt; 18 years of age)</td>
<td>$ 620.50</td>
<td>MBS item number 12213</td>
</tr>
<tr>
<td>Diagnostic sleep study (&gt; 18 years of age)</td>
<td>$ 577.05</td>
<td>MBS item number 12203</td>
</tr>
<tr>
<td>Ventilation titration sleep study (&lt; 18 years of age)</td>
<td>$ 620.50</td>
<td>MBS item number 12217</td>
</tr>
<tr>
<td>Ventilation titration sleep study (&gt; 18 years of age)</td>
<td>$ 577.05</td>
<td>MBS item number 12207</td>
</tr>
<tr>
<td><strong>Ventilator provision, maintenance and support</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resmed VPAP IV bi-level ventilator</td>
<td>$ 2950.00</td>
<td>CPAP Direct</td>
</tr>
<tr>
<td>Resmed power station II (backup power supply)</td>
<td>$ 577.00</td>
<td>CPAP Australia</td>
</tr>
<tr>
<td>Onsite ventilator maintenance and repair (per annum)</td>
<td>$ 295.00</td>
<td>(NICE, 2010)</td>
</tr>
<tr>
<td>Outreach nurse, per visit</td>
<td>$ 75.00</td>
<td>Researcher estimate</td>
</tr>
<tr>
<td><strong>Consumables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resmed Mirage FX nasal mask and headset</td>
<td>$ 239.00</td>
<td>CPAP Direct</td>
</tr>
<tr>
<td>Ventilator tubing (generic, 1.8 meters)</td>
<td>$ 19.50</td>
<td>CPAP Australia</td>
</tr>
<tr>
<td>Ventilator filters (pack of 3)</td>
<td>$ 7.50</td>
<td>CPAP Australia</td>
</tr>
</tbody>
</table>
Model parameters

The following tables (Tables 27, 28 & 29) list: (i) the model parameters that were used in the base-case analysis; (ii) the parameter names that were used to represent them in the decision-analytic model; (iii) the ranges that were examined in uncertainty and sensitivity analysis and; (iv) the data sources from which this information was obtained. The first of these three tables, Table 27, contains a list of: (i) the proportion of deaths in the cohorts that are due to severe cardiomyopathy (SCM_deaths); (ii) the proportion of the cohorts in Intervention 1 and 2 with severe cardiomyopathy who were ventilated prior to death (IntV1_SCM_V & IntV2_SCM_V) and; (iii) the condition specific and generic utility weighted health-related quality of life of individuals, prior and after the initiation of LTMV (MRF28_NoV; AQoL_NoV & MRF28_V; AQoL_V respectively). The estimate of the proportion of the cohort with severe cardiomyopathy who were ventilated prior to death in Intervention 1 and 2 was derived from the mean age of death of 16.9 years as reported in Eagle et al (2002).

In a decision-analytic model, unlike in a Markov model, there is no in-built mechanism for accounting for the passage of time. This must be explicitly defined so that outcomes such as quality adjusted survival can take into account changes in the health status of individuals over time, and so that discounting can be applied to costs and outcomes at the time they were incurred (Drummond et al, 2005). The second of the three tables, Table 28, lists the life-years for the members of the three cohorts that were lived without the use of LTMV to support respiratory function (YnoL=years not ventilated) and the life-years that were lived with LTMV (YV=years ventilated). The values of these variables and the ranges that were examined in the sensitivity analysis were derived from the age at which LTMV was initiated in the two interventions and the current practice comparator, and the age of death (by primary cause) as reported in Eagle et al (2002).

The third table, Table 29, lists the mean undiscounted total cost for 100 individuals for each of the terminal nodes in the model and the ranges that were examined in the sensitivity analysis. The following method was used to generate the cost estimates: (i) the activities in the event pathways were entered into the rows of an excel spreadsheet, along with the frequency of their occurrence; (ii) the unit cost for the activities were entered into the columns directly to the right of these rows; (ii) the age of the individuals, in six monthly increments (i.e. 15.0, 15.5 years) were entered a series of columns to the right of the activity data; (iii) the cost of each activity was entered into the column that corresponded to the point of time at which the cost was incurred, based on the frequency of their occurrence (i.e. in the current practice pathway, an individual has diagnostic sleep study once a year from 15.0 years of age until the initiation of LTMV at 18.5 years of age and from this point on they a six monthly titration study); (iv) the total costs for individuals with and without severe cardiomyopathy, prior to and after the initiation of ventilation (were
relevant) were summed and this was then divided by period of time over which the costs were accumulated to return a mean annual cost; (v) the mean annual cost was multiplied by 100 to calculate the cost for 100 individuals and; (vi) plus and minus 25% of the total cost were calculated for use in the sensitivity analysis.
Table 27: Proportions of death by cause, proportion ventilated by cause and HRQoL

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Description</th>
<th>Base-case value</th>
<th>Range for sensitivity analysis</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proportion of deaths, by primary cause</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCM_deaths</td>
<td>Deaths due to severe cardiomyopathy</td>
<td>0.1475</td>
<td>0.0995 - 0.2074</td>
<td>Systematic review&lt;sup&gt;80&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Proportion of cohort with SCM ventilated, by pathway&lt;sup&gt;81&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IntV1_SCM_V</td>
<td>Intervention 1, severe cardiomyopathy</td>
<td>1.000</td>
<td>Not applicable</td>
<td>Derived</td>
</tr>
<tr>
<td>IntV2_SCM_V</td>
<td>Intervention 2, severe cardiomyopathy</td>
<td>0.000</td>
<td>Not applicable</td>
<td>Derived</td>
</tr>
<tr>
<td><strong>Health related quality of life (HRQoL)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRF28_NoV</td>
<td>Condition specific HRQoL&lt;sup&gt;82&lt;/sup&gt;, without LTMV</td>
<td>0.8660</td>
<td>0.4213 - 0.9964</td>
<td>Small scale study</td>
</tr>
<tr>
<td>MRF28_V</td>
<td>Condition specific HRQoL, with LTMV</td>
<td>0.6165</td>
<td>0.5995 - 0.6250</td>
<td>Small scale study</td>
</tr>
<tr>
<td>AQoL_NoV</td>
<td>Utility weighted quality of life, without LTMV</td>
<td>0.1349</td>
<td>0.0982 - 0.1716</td>
<td>Small scale study</td>
</tr>
<tr>
<td>AQoL_V</td>
<td>Utility weighted quality of life, with LTMV</td>
<td>0.1727</td>
<td>0.0927 - 0.2527</td>
<td>Small scale study</td>
</tr>
</tbody>
</table>

<sup>80</sup> Based on the findings of the study of Eagle et al (2002)

<sup>81</sup> In the current practice comparator, there were no individuals with severe cardiomyopathy who lived until 18.5 years of age, the point at which LTMV was initiated

<sup>82</sup> The scores of the MRF28 were inverted (subtracted from 1) so that higher scores were associated with a better quality of life, and lower scores with a poorer quality of life so that the scores could be used to generate incremental Cost-effectiveness ratios.
Table 28: Life years spent without and with LTMV

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Description</th>
<th>Base-case value</th>
<th>Range for sensitivity analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP_SCM_YnoV</td>
<td>Current practice SCM, life years without LTMV</td>
<td>1.90</td>
<td>0.23 - 2.97</td>
</tr>
<tr>
<td>CP_NoSCM_YnoV</td>
<td>Current practice no SCM, life years without LTMV</td>
<td>3.50</td>
<td>Not applicable</td>
</tr>
<tr>
<td>CP_NoSCM_YV</td>
<td>Current practice no SCM, life years with LTMV</td>
<td>6.80</td>
<td>4.61 - 8.08</td>
</tr>
<tr>
<td>IntV1_SCM_YnoV</td>
<td>Intervention 1 SCM, life years without LTMV</td>
<td>0.50</td>
<td>0.23 - 0.50</td>
</tr>
<tr>
<td>IntV1_SCM_YV</td>
<td>Intervention 1 SCM, life years with LTMV</td>
<td>1.40</td>
<td>0.00 - 2.47</td>
</tr>
<tr>
<td>IntV1_NoSCM_YNoV</td>
<td>Intervention 1 no SCM, life years without LTMV</td>
<td>0.50</td>
<td>Not applicable</td>
</tr>
<tr>
<td>IntV1_NoSCM_YV</td>
<td>Intervention 1 no SCM, life years with LTMV</td>
<td>9.80</td>
<td>7.61 - 11.08</td>
</tr>
<tr>
<td>IntV2_SCM_YnoV</td>
<td>Intervention 2 SCM, life years without LTMV</td>
<td>1.90</td>
<td>0.23 - 2.50</td>
</tr>
<tr>
<td>IntV2_SCM_YV</td>
<td>Intervention 2 SCM, life years with LTMV</td>
<td>0.00</td>
<td>0.00 – 0.47</td>
</tr>
<tr>
<td>IntV2_NoSCM_YNoV</td>
<td>Intervention 2, no SCM, life years without LTMV</td>
<td>2.50</td>
<td>Not applicable</td>
</tr>
<tr>
<td>IntV2_NoSCM_YV</td>
<td>Intervention 2, no SCM, life years with LTMV</td>
<td>7.80</td>
<td>5.61 - 9.08</td>
</tr>
</tbody>
</table>
# Table 29: Mean annual cost for 100 individuals (undiscounted)

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Description</th>
<th>Base-case value</th>
<th>Range for sensitivity analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current practice</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP_SCM_NoVP</td>
<td>Severe cardiomyopathy (standard care)</td>
<td>$ 188,133</td>
<td>$ 141,100 - $ 235,167</td>
</tr>
<tr>
<td>CP_NoSCM_NoVP</td>
<td>Cardiomyopathy not severe (standard care)</td>
<td>$ 149,029</td>
<td>$ 111,771 - $ 186,286</td>
</tr>
<tr>
<td>CP_NoSCM_VP</td>
<td>Cardiomyopathy not severe (standard care &amp; LTMV)</td>
<td>$ 418,985</td>
<td>$ 314,238 - $ 523,731</td>
</tr>
<tr>
<td><strong>Intervention 1: SDB in REM without hypercapnia</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IntV1_SCM_NoVP</td>
<td>Severe cardiomyopathy (standard care)</td>
<td>$ 102,700</td>
<td>$ 77,025 - $ 128,375</td>
</tr>
<tr>
<td>IntV1_SCM_VP</td>
<td>Severe cardiomyopathy (standard care &amp; LTMV)</td>
<td>$ 737,200</td>
<td>$ 552,900 - $ 921,500</td>
</tr>
<tr>
<td>IntV1_NoSCM_NoVP</td>
<td>Cardiomyopathy not severe (standard care)</td>
<td>$ 102,700</td>
<td>$ 77,025 - $ 128,375</td>
</tr>
<tr>
<td>IntV1_NoSCM_VP</td>
<td>Cardiomyopathy not severe (standard care &amp; LTMV)</td>
<td>$ 393,074</td>
<td>$ 294,805 - $ 491,342</td>
</tr>
<tr>
<td><strong>Intervention 2: SDB in REM without hypercapnia</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IntV2_SCM_NoVP</td>
<td>Severe cardiomyopathy (standard care)</td>
<td>$ 188,133</td>
<td>$ 141,100 - $ 235,167</td>
</tr>
<tr>
<td>IntV2_SCM_VP</td>
<td>Severe cardiomyopathy (standard care &amp; LTMV)</td>
<td>$ 431,100</td>
<td>$ 323,325 - $ 538,875</td>
</tr>
<tr>
<td>IntV2_NoSCM_NoVP</td>
<td>Cardiomyopathy not severe (standard care)</td>
<td>$ 153,960</td>
<td>$ 115,470 - $ 192,450</td>
</tr>
<tr>
<td>IntV2_NoSCM_VP</td>
<td>Cardiomyopathy not severe (standard care &amp; LTMV)</td>
<td>$ 393,074</td>
<td>$ 309,110 - $ 515,183</td>
</tr>
</tbody>
</table>
The following formulas were used in Treeage Pro 2009 to calculate the costs and outcomes for each of the terminal nodes of the model (with Intervention 1 used as example)

**Cardiomyopathy severe, not ventilated prior to death**

The life-years spent without ventilation for individuals who have severe cardiomyopathy who die prior the initiation of LTMV was used to calculate the (undiscounted and discounted) total cost (a), condition specific HRQoL (b) and generic utility weighted HRQoL (c) for terminal nodes 1, 4 & 7.

(a) \[\text{Discount}(\text{IntV1}_Y\text{noV}\times\text{IntV1}_Y\text{noV};\text{Disc}_\text{Rate})\]
(b) \[\text{Discount}(100\times\text{IntV1}_Y\text{noV}\times\text{MRF28}_\text{noV};\text{IntV1}_Y\text{noV};\text{Disc}_\text{Rate})\]
(c) \[\text{Discount}(100\times\text{IntV1}_Y\text{noV}\times\text{AQoL}_\text{noV};\text{IntV1}_Y\text{noV};\text{Disc}_\text{Rate})\]

**Cardiomyopathy, severe, ventilated prior to death**

The life-years spent without ventilation, and with ventilation for individuals who have severe cardiomyopathy was used to calculate the (undiscounted and discounted) total cost (a), condition specific HRQoL (b) and generic utility weighted HRQoL (c) for terminal nodes 2 & 5.

(a) \[\text{Discount}((\text{IntV1}_\text{YnoV}\times\text{IntV1}_\text{YnoV});\text{IntV1}_Y\text{noV};\text{Disc}_\text{Rate})+\text{Discount}((\text{IntV1}_\text{YV}\times\text{IntV1}_\text{YV};\text{Disc}_\text{Rate}))\]
(b) \[\text{Discount}(100\times(\text{IntV1}_\text{YnoV}\times\text{MRF28}_\text{noV};\text{IntV1}_Y\text{noV};\text{Disc}_\text{Rate})+\text{Discount}(100\times(\text{IntV1}_\text{YV}\times\text{MRF28}_\text{V};\text{IntV1}_Y\text{V};\text{Disc}_\text{Rate}))\]
(c) \[\text{Discount}(100\times(\text{IntV1}_\text{YnoV}\times\text{AQoL}_\text{noV};\text{IntV1}_Y\text{noV};\text{Disc}_\text{Rate})+\text{Discount}(100\times(\text{IntV1}_\text{YV}\times\text{AQoL}_\text{V};\text{IntV1}_Y\text{V};\text{Disc}_\text{Rate}))\]

**Cardiomyopathy, not severe, ventilated**

The life-years spent without ventilation, and with ventilation for individuals who did not have severe cardiomyopathy was used to calculate the (undiscounted and discounted) total cost (a), condition specific HRQoL (b) and generic utility weighted HRQoL (c) for terminal nodes 3, 6 & 8.

(a) \[\text{Discount}((\text{IntV1}_\text{YnoV}\times\text{IntV1}_Y\text{noV};\text{IntV1}_\text{YnoV};\text{Disc}_\text{Rate})+\text{Discount}((\text{IntV1}_\text{YV}\times\text{IntV1}_Y\text{V};\text{Disc}_\text{Rate}))\]
(b) \[\text{Discount}(100\times(\text{IntV1}_\text{YnoV}\times\text{MRF28}_\text{noV};\text{IntV1}_Y\text{noV};\text{Disc}_\text{Rate})+\text{Discount}(100\times(\text{IntV1}_\text{YV}\times\text{MRF28}_\text{V};\text{IntV1}_Y\text{V};\text{Disc}_\text{Rate}))\]

---

83 A discount rate of 0.00
Model validation

The descriptive validity of the decision analytic model was evaluated through an examination of its structure, data sources, problem formulation and results. The model is a relatively simple representation of the key events in the natural history of the disease and the use of LTMV to support respiratory function. While it is recommended practice that a model be no more complex than is absolutely necessary to address the decision problem, the structure and complexity of the decision-analytic model was limited by the availability of data. In the published literature the disease is characterised as having two main causes of death, these being primarily cardiac in nature (at mean age of 16.9 years) and primarily respiratory in nature (at a mean age of 25.3 years when LTMV is used). The two causes of death are primarily used in the model to determine the proportion of individuals with severe cardiomyopathy who are ventilated prior to their death, and the cost of doing so, when the age of initiation of LTMV is reduced from 18.5 years (in the Current practice comparator) to 17.5 years (in Intervention 2) and 15.5 years (in Intervention 1). While this may not reflect the complexity of clinical practice, where individuals in chronic respiratory failure also have progressive cardiomyopathy, the data needed to account for this in the model was not available.

The following sources of data were used to structure and populate the model. The evidence for the effectiveness of the intervention was obtained from a formal systematic literature review that was conducted by the author. The evidence base of the intervention was found to extremely limited and weak, with the decision analytic model ultimately being based on the findings of a single study, that of Eagle et al (2002). The impact of LTMV on health related quality of life was obtained from a small scale study that was conducted by the author, which was unable to recruit and retain a large enough sample to be able to detect a statistically significant effect. The treatment pathways for the clinical management of DMD in the model were based on best practice guidelines, which were developed by specialists in the United Kingdom, Europe and the United States. As there were no published estimates of the cost of delivering domiciliary LTMV for individuals with DMD, a service delivery model was developed by the author to provide a basis for a cost-estimate. There was no use of expert opinion as a source of data.

The problem that was addressed in the decision analytic model was the timing of the initiation of an intervention to address the primary cause of death in DMD. The outcomes in the model, survival and health related quality of life, are important patient relevant outcomes. The setting in which the intervention

---

84 Discount rate (mean of 0.03, range of 0.00 to 0.10)
was used, at home via an outpatient service delivery model is believed to reflect how the intervention is initiated and delivered in most states and territories in Australia. The target population of the intervention were individuals with DMD who were at risk of developing sleep disordered breathing, as a consequence of a decline in their muscle strength and function following the loss of ambulation. The time horizon of the analysis was the life-time of the target population to enable the evaluation to capture all of the costs and effects of the intervention.

The descriptive validity of a model should ideally be assessed by external peer review by a group of disinterested but knowledgeable experts. Duchenne muscular dystrophy is a rare disease and respiratory function in neuromuscular disease and the use of long-term mechanical ventilation is a highly specialised area of knowledge and consequently while there are knowledgeable experts, they are not likely to be disinterested observers. There were two members of the supervisory panel of the thesis who were senior neurologists at two of the tertiary referral hospitals in Melbourne who were directly or indirectly responsible for the clinical management of individuals with DMD. Therefore, a judgement will need to be made by the end users of the evaluation about whether or not the model has descriptive validity.

The steps that were undertaken to evaluate the internal validity of the model were as follows: (i) the parameter values were carefully checked against the data in Tables 27, 28 & 29; (ii) the results produced by the formulas in the model were checked against manual calculations; (iii) the parameter values were varied one at a time from their lowest to highest possible values (from the 95% CIs or ranges for the parameters as listed in Tables 27, 28 & 29) to determine whether the resultant changes in the model outcomes were as expected and; (iv) the impact of two extreme scenarios were examined. In the first of the two extreme scenarios all of the model parameters (other than the discount rate) were set to their lowest values and in the second scenario; they were all set to their highest values (from the 95% CIs & ranges listed in tables 27, 28 & 29).

There were no unexpected changes in model outcomes as a result of the changes made to single parameter values. The following examples (all costs and outcomes discounted by 3%) are provided to help illustrate this:\ref{footnote2}; (i) at the low end of the 95% CI for pre-ventilation utility weighted quality of life (AQLo1_NoV) there was a decrease in the effectiveness of the comparator (a total of 124.80 QALYs) and of Intervention 1 and 2 (totals of 130.41 & 142.67 QALYs respectively); (ii) at the high end of the 95% CI for pre-ventilation utility weighted quality of life (AQLo1_NoV) there was a slight increase in the effectiveness of the comparator (a total of 147.73 QALYs) and of Intervention 1 (a total of 146.29 QALYs)

\footnote{2 In the base-case Cost-utility analysis the comparator produced 136.27 QALYs and intervention 1 and 2 produced 144.48 and 138.94 QALYs respectively.}
and Intervention 2 (a total of 147.46 QALYs); (iii) at the low end of the 95% CI for the number of life-years following the initiation of LTMV in Intervention 1, for individuals who did not have SCM (IntV1_NoSCM_YV), there a decrease in the effectiveness of Intervention 1 (a total of 115 QALYs) but no change in the effectiveness of the comparator and intervention 2 and; (iv) at the high end of the 95% CI for the number of life-years following the initiation of LTMV in Intervention 1, for individuals who did not have SCM (IntV1_NoSCM_YV), there was increase in the effectiveness of Intervention 1 (a total of 161.5 QALYs) and there was no change in the effectiveness of the comparator and intervention 2.

The two extreme scenarios did not produce any unexpected changes in model outcomes. In the first of the two scenarios, in which all the variables were set to their lowest values, there was a fall in the cost of the two interventions and the comparator. The cost of the comparator fell to $1.2 million from its base-case value of $2.9 million and the cost of Intervention 1 and 2 fell to $1.8 million and $1.9 million respectively. There was a decrease in the number of QALYs produced by the comparator and the two interventions and an increase in the incremental effectiveness of the two interventions, with Intervention 1 producing 67.3 QALYs and Intervention 2 producing 60 QALYs. This increase was the result of a reduction in the number of individuals who were ventilated and the length of time that they were ventilated.

In the second of the two extreme scenarios, there was a relatively minor but uniform increase in the total lifetime cost of the comparator and intervention 1 and 2 (to $3.8, $4.0 & $4.5 million respectively). There was also a uniform increase in the effectiveness of the comparator and intervention 1 and 2, which produced 207.1, 213.4 and 226.9 QALYs respectively. The increase in cost was the result of an increase in the number of individuals who were ventilated and an increase in the period of time over which they were ventilated. The increase in effectiveness was the result of a minor increase in post-ventilation utility weighted HRQoL of 8.11% , as compared to a 3.78% post-ventilation increase in the base-case analysis.

The external validity of the model was evaluated through a comparison to the findings of the study of Eagle et al (2002). The decision analytic model was based on data obtained from this study, specifically:

(i) the proportion of deaths due to severe cardiomyopathy (as a primary cause); (ii) the proportion of deaths due to respiratory failure (as a primary cause); (iii) the mean age of death from severe cardiomyopathy (as a primary cause) and; (iii) the mean age of death due to respiratory failure (as a primary cause). The Monte Carlo simulation (probabilistic sensitivity analysis) function in Treeage Pro 2009 produces summary statistics, for each simulation run, for all of the model parameters with defined distributions. The summary statistics were used to establish in a dependent validation (Caro et al, 2012; Eddy et al, 2012) that the mean, minimum and maximum values of the above listed model parameters (for

---

86 The discount rate was kept at 3% to simplify the comparison
a simulation run of 10,000) were in agreement with the range of values that were reported by Eagle et al (2002).

The predictive validity of the model was not able to be evaluated due to its use of long-term as opposed to short-term outcome measures, and the lack of clinical trial or a published research design for a study to collect the data required to this.
Conclusion

An exploratory Cost-effectiveness and Cost-utility analysis was conducted from a health sector perspective of the lifetime costs and outcomes for two alternative treatment pathways for the use of domiciliary non-invasive positive pressure LTMV in DMD. A small scale study was undertaken as part as the evaluation to collect health related quality of life data for use in a decision-analytic model. Unfortunately only 12 of an estimated total of 82 potential study participants agreed to participate in the study of which only 11 individuals returned their study questionnaires. Consequently, the small scale study was unable to recruit the minimum number of participants that was needed to be able to detect a statistically significant difference in the HRQoL of ventilated and unventilated individuals.

Economic evaluations are conducted to inform timely decision making about the use of scarce health care resources under conditions of uncertainty, and it is not unusual for such decisions to be made when the best available evidence is weak and surrounded by significant uncertainty. It is also not unusual for a decision-analytic model to include data from a range of sources of varying quality, and to not exclude the use of data solely on the basis that it was not possible to demonstrate a statistically significant difference between two groups. While these models may lack in scientific rigor, they provide a basis through which the best available evidence can be structured in a systematic and logical manner to better inform timely decision making.

The next chapter of the thesis describes in detail the results of the base-case analysis and the findings of simple, multivariate and probabilistic sensitivity analysis.
Chapter 8: Results

Introduction

The previous chapter of the thesis described in detail the methods that were used in the economic evaluation to combine, in a decision analytic model; the costs of the events in the clinical treatment pathways, the evidence for the effectiveness of the intervention and the results of the small scale HRQoL study. The role of this chapter in the thesis is to describe in detail the results of the base-case analysis of the economic evaluation and the findings of uncertainty and sensitivity analysis.

The chapter begins with the results of the base-case Cost-effectiveness and Cost-utility analyses. This is followed by the methods that were used in the univariate, multivariate and probabilistic sensitivity analyses and the findings of these analyses.
Base-case analysis

Cost-effectiveness analysis

In the base-case of the Cost-effectiveness analysis (all costs and outcomes discounted by 3%) the initiation of LTMV from the onset of SDB in REM with hypercapnia (Intervention 2) produced a total of 585.31 life-years, at a cost of $4,811 a year. The initiation of LTMV from the onset of SDB in REM without hypercapnia (Intervention 1) produced a total of 534.75 life-years, at a cost of $6,090 a year. The two interventions were dominated by the current practice comparator, which produced a total of 606.53 life years, at a cost of $4,550 a year. The results are presented in graphical form in Figure 6 (all costs and outcomes were discounted at 3%).

The following table, Table 30, contains the details of the results of the Cost-effectiveness analysis. Starting from the left most column, the table contains: (i) the name of intervention or comparator; (ii) the total cost; (iii) the incremental cost (where applicable); (iv) the effectiveness (condition specific quality adjusted life expectancy calculated from results for the MRF28); (v) the incremental effectiveness (vi); the Cost-effectiveness and; (vii) the incremental Cost-effectiveness ratio (where applicable). The results are reported as undiscounted and discounted at 3% and 5% per annum.
<table>
<thead>
<tr>
<th>Intervention</th>
<th>Cost</th>
<th>Incremental Cost</th>
<th>Effectiveness (MRF28)</th>
<th>Incremental Effectiveness</th>
<th>Cost/Effectiveness (C/E)</th>
<th>Incremental C/E (ICER)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Undiscounted</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current practice</td>
<td>$2,926,246</td>
<td>Not applicable</td>
<td>640.05</td>
<td>Not applicable</td>
<td>$4,572 a life year</td>
<td>Not applicable</td>
</tr>
<tr>
<td>SBD in REM, with hypercapnia (Intervention 2)</td>
<td>$2,994,597</td>
<td>$68,351</td>
<td>618.78</td>
<td>-21.27</td>
<td>$4,840 a life year</td>
<td>Dominated</td>
</tr>
<tr>
<td>SBD in REM, without hypercapnia (Intervention 1)</td>
<td>$3,487,519</td>
<td>$492,519</td>
<td>571.09</td>
<td>-68.96</td>
<td>$6,107 a life year</td>
<td>Dominated</td>
</tr>
<tr>
<td><strong>Discounted (3% per annum)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current practice</td>
<td>$2,759,813</td>
<td>Not applicable</td>
<td>606.53</td>
<td>Not applicable</td>
<td>$4,550 a life year</td>
<td>Not applicable</td>
</tr>
<tr>
<td>SBD in REM, with hypercapnia (Intervention 2)</td>
<td>$2,815,753</td>
<td>$55,940</td>
<td>585.31</td>
<td>-21.21</td>
<td>$4,811 a life year</td>
<td>Dominated</td>
</tr>
<tr>
<td>SBD in REM, without hypercapnia (Intervention 1)</td>
<td>$3,256,695</td>
<td>$440,942</td>
<td>534.75</td>
<td>-71.78</td>
<td>$6,090 a life year</td>
<td>Dominated</td>
</tr>
<tr>
<td><strong>Discounted (5% per annum)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current practice</td>
<td>$2,654,223</td>
<td>Not applicable</td>
<td>585.19</td>
<td>Not applicable</td>
<td>$4,536 a life year</td>
<td>Not applicable</td>
</tr>
<tr>
<td>SBD in REM, with hypercapnia (Intervention 2)</td>
<td>$2,702,636</td>
<td>$48,413</td>
<td>564.08</td>
<td>-21.11</td>
<td>$4,791 a life year</td>
<td>Dominated</td>
</tr>
<tr>
<td>SBD in REM, without hypercapnia (Intervention 1)</td>
<td>$3,111,602</td>
<td>$408,966</td>
<td>511.90</td>
<td>-73.29</td>
<td>$6,079 a life year</td>
<td>Dominated</td>
</tr>
</tbody>
</table>
In the above figure (Figure 6) the current practice comparator (represented by a triangle) is dominant, as it is both more effective and less costly than Intervention 1 & 2.
Cost-utility analysis

In the base-case of the Cost-utility analysis (all costs and outcomes discounted by 3%) the initiation of LTMV from the onset of SDB in REM with hypercapnia (Intervention 2) produced a total of 138.94 QALYs at a cost of $ 20,267 a QALY. The initiation of LTMV from the onset of SDB in REM without hypercapnia (Intervention 1) produced a total of 144.38 QALYs at a cost of $ 22,541 a QALY, with the current practice comparator producing a total of 136.27 QALYs at a cost of $ 20,253 a QALY. The two interventions produced only a small increase in the number of QALYs above that of the comparator of 5.54 and 2.67 QALYs (Intervention 1 & 2 respectively) and there were no strategies that were clearly dominated by any other, in either standard or extended dominance. The results are presented in graphical form in Figure 7 (all costs and outcomes discounted at 3%).

The following table, Table 31, contains the details of the results of the Cost-utility analysis. Starting from the left most column, the table contains: (i) the name of intervention or comparator; (ii) the total cost; (iii) the incremental cost (where applicable); (iv) the effectiveness (utility weighted quality adjusted life-years, calculated from the results of the AQoL); (v) the incremental effectiveness (vi); the Cost-utility and; (vii) the incremental Cost-utility ratio (where applicable). The results are reported as undiscounted and discounted at 3% and 5% per annum.
Table 31: Summary of results - Cost-utility analysis (QALYs)

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Effectiveness (AQoL1)</th>
<th>Incremental Effectiveness</th>
<th>Cost/Effectiveness (C/E)</th>
<th>Incremental C/E (ICER)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Undiscounted</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current practice</td>
<td>144.15</td>
<td>Not applicable</td>
<td>$ 20,301 QALY</td>
<td>Not applicable</td>
</tr>
<tr>
<td>SBD in REM, with hypercapnia (Intervention 2)</td>
<td>147.37</td>
<td>3.22 QALYs</td>
<td>$ 20,321 QALY</td>
<td>$ 21,211</td>
</tr>
<tr>
<td>SBD in REM, without hypercapnia (Intervention 1)</td>
<td>154.59</td>
<td>7.23 QALYs</td>
<td>$ 22,559 QALY</td>
<td>$ 68,220</td>
</tr>
<tr>
<td><strong>Discounted (3% per annum)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current practice</td>
<td>136.27</td>
<td>Not applicable</td>
<td>$ 20,253 QALY</td>
<td>Not applicable</td>
</tr>
<tr>
<td>SBD in REM, with hypercapnia (Intervention 2)</td>
<td>138.94</td>
<td>2.67 QALYs</td>
<td>$ 20,267 QALY</td>
<td>$ 20,968</td>
</tr>
<tr>
<td>SBD in REM, without hypercapnia (Intervention 1)</td>
<td>144.48</td>
<td>5.54 QALYs</td>
<td>$ 22,541 QALY</td>
<td>$ 79,548</td>
</tr>
<tr>
<td><strong>Discounted (5% per annum)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current practice</td>
<td>131.26</td>
<td>Not applicable</td>
<td>$ 20,221 QALY</td>
<td>Not applicable</td>
</tr>
<tr>
<td>SBD in REM, with hypercapnia (Intervention 2)</td>
<td>133.59</td>
<td>2.33 QALYs</td>
<td>$ 20,230 QALY</td>
<td>$ 20,755</td>
</tr>
<tr>
<td>SBD in REM, without hypercapnia (Intervention 1)</td>
<td>138.12</td>
<td>4.53 QALYs</td>
<td>$ 22,528 QALY</td>
<td>$ 90,345</td>
</tr>
</tbody>
</table>
In the above figure (Figure 7), the two interventions are more effective than the current practice comparator but they are also more costly.
Uncertainty and sensitivity analysis

The results of a model are conditional on the parameter data and assumptions that are used in its construction and consequently it should not be seen as providing a fixed, accurate point estimate of the Cost-effectiveness or cost-utility of an intervention. As previously discussed, it is recommended practice in economic evaluations to use univariate and multivariate sensitivity analysis to examine the impact of the decisions that were made in relation to the structure of a model, and probabilistic sensitivity analysis to examine parameter and overall uncertainty (Weinstein et al, 2003; Briggs, 2005; Andronis, Barton & Bryan, 2009; Briggs et al, 2012; Caro et al, 2012).

Univariate sensitivity analysis

A total of 24 of the 29 model parameters listed in Tables 27, 28 and 29 (and the discount rate) were included in the univariate sensitivity analysis. The five model parameters that were excluded were those with fixed values and no associated 95% confidence interval or range. The five parameters were used in the model to represent the fixed period of time between an individual’s entry into the model (for individuals without severe cardiomyopathy) and the initiation of LTMV (CP_NoSCM_YV, IntV1_NoSCM_YNoV & IntV2_NoSCM_YNoV) and the proportion of individuals (with severe cardiomyopathy) in Interventions 1 & 2 who were ventilated prior to the death from severe cardiomyopathy (IntV1_SCM_V & IntV2_SCM_V).

The univariate sensitivity analysis was undertaken through the use of the one-way sensitivity analysis function in Treeage Pro 2009. This function recalculated the results of the evaluation across the range of the possible values for the model parameters as listed in Tables 27, 28 & 29. For the purpose of the analysis: (i) a small change in the value of a parameter was defined as an increase or decrease of 20% (or less) of the value used in the base-case analysis and; (ii) the results were seen to be sensitive to the model parameter, if there was a 50% increase or decrease (above or below the result of the base-case analysis) in an incremental Cost-effectiveness or Cost-utility ratio.
Cost-effectiveness analysis

The results of the Cost-effectiveness analysis were found to be sensitive to small changes in the following model parameters: (i) the number of years of life after the initiation of LTMV (for individuals without SCM) in the current practice comparator (CP_NoSCM_YV) and; (ii) the number of life years after the initiation of LTMV (for individuals without SCM) in Intervention 2 (IntV2_NoSCM_YV). The results were found not to be sensitive to changes in the discount rate.

The results of the univariate sensitivity analysis are displayed graphically in a tornado diagram\(^87\) in Figure 8. The parameters in the figure are labelled with the names that were used in the model (as listed above) along with their upper and lower ranges (as listed in Table 27). The dotted vertical line in the diagram is equal to the base-case point estimate of approximately $4,500 a life year.

**Figure 8: Univariate analysis - Tornado diagram (Life years)**

---

\(^87\) In a tornado diagram: the outcomes are represented on the horizontal axis; the parameters are arranged on the vertical axis; the width of each bar corresponds to the outcome range for the specified parameter; the parameter that is associated with the greatest uncertainty is placed at the top and; a vertical dotted line passes through the bars that corresponds to the value of result of the base-case analysis (i.e. per life year or QALY) (Briggs et al, 2012).
It is clearly evident in the Tornado diagram, in Figure 8, that the greatest area of uncertainty in the Cost-effectiveness analysis was the number of years of life after the initiation of LTMV (for individuals without SCM) in the current practice comparator (CP_NoSCM_YV). There was much less uncertainty associated with the number of years of life after the initiation of LTMV (for individuals without SCM) in Intervention 2 (IntV2_NoSCM_YV).

A threshold analysis was conducted in Treeage Pro 2009, through the use of the program’s threshold analysis function, to quantify the impact of the two parameters on the results of the Cost-effectiveness analysis. A maximum acceptable ratio of $50,000 a life year was used in the threshold analysis with the results reported in terms of the net health benefit.

The results of the analysis were as follows:

1. **CP_NoSCM_YV**: When there were 4.61 to 6.26 post-ventilation years of life (for individuals without SCM) in the Current practice comparator, Intervention 2 was dominant (threshold value of 529.0 life years) and; when there were 6.26 years or greater, the Current practice comparator was dominant.
2. **IntV2_NoSCM_YV**: When there were 5.61 to 8.33 post-ventilation years of life (for individuals without SCM) in Intervention 2, the Current practice comparator was dominant (threshold value of 551.3 life years) and; when there were 8.33 years or greater, Intervention 2 was dominant.

---

88 A threshold analysis is a form of univariate sensitivity analysis in which the value of a model parameter is varied to determine the point at which it produces a major change in the results of an analysis (Walker & Miners, 2005; Manning, Fryback & Weinstein, 1996).
Cost-utility analysis

The results of the Cost-utility analysis were found to be sensitive to small changes in the following model parameters: (i) pre & post ventilation utility weighted HRQoL (AQoL_NoV & AQoL_V); (ii) the cost of treatment prior to the initiation of ventilation (for individuals without SCM) in the current practice comparator and intervention 2 (CP_NoSCM_NoVP & IntV2_NoSCM_NoVP); (iii) the cost of treatment after the initiation of ventilation (for individuals without SCM) in the current practice comparator and interventions 1 and 2 (CP_NoSCM_VP; IntV1_NoSCM_VP & IntV2_NoSCM_VP); and; (iv) the number of years of life after the initiation of ventilation (for individuals without SCM) in the current comparator practice and interventions 1 & 2 (CP_NoSCM_YV; IntV1_NoSCM_YV; IntV2_NoSCM_YV). The results were found not to be sensitive to changes in the discount rate.

The results of the univariate sensitivity analysis of the Cost-utility analysis are displayed graphically in a tornado diagram in Figure 9. The parameters are labelled in the diagram with the names that were used in the model (as listed above) along with their upper and lower ranges (as listed in Tables 27, 28 & 29). The dotted vertical line in the diagram is equal to the base-case point estimate of approximately $21,000 a QALY.

Figure 9: Univariate analysis - Tornado diagram (QALYs)

Tornado Diagram: CUA
Timing of the initiation of LTMV

![Tornado Diagram](image-url)
It is clearly evident in the Tornado diagram, in Figure 9, that the greatest areas of uncertainty in the Cost-utility analysis (in decreasing order of importance) were: (i) post ventilation utility weighted quality of life (AQoL_V); (ii) the post ventilation annual cost of treatment (for individuals without SCM) in the Current practice comparator (CP_NoSCM_VP); (iii) the post ventilation annual cost of treatment (for individuals without SCM) in Intervention 2 (IntV2_NoSCM_VP); (iv) the post ventilation annual cost of treatment (for individuals without SCM) in Intervention 1 (IntV1_NoSCM_VP) and; (v) pre ventilation utility weighted quality of life (AQoL_NoV).

A threshold analysis was conducted in Treeage Pro 2009 to quantify the impact of these parameters had on the results of the Cost-utility analysis. A maximum acceptable ratio of $50,000 a QALY was used in the threshold analysis, with the results reported in terms of the net health benefit.

The results of the analysis were as follows:

1. **AQoL_NOV**: When pre-ventilation utility weighted HRQoL was 0.10 to 0.12, Intervention 1 was dominant (threshold value of 78.4 QALYs); when it was 0.12 to 0.15, Intervention 2 was dominant (threshold value of 87.1 QALYs) and; when it 0.15 or greater, the Current practice comparator was dominant;

2. **AQoL_V**: When post-ventilation utility weighted HRQoL was 0.09 to 0.15, the Current practice comparator was dominant (threshold value of 70.2 QALYs); when it was 0.15 to 0.19, Intervention 2 was dominant (threshold value of 94.2 QALYs) and; when it was 0.19 or greater, Intervention 1 was dominant;

3. **CP_NoSCM_NoVP**: When the pre-ventilation cost of treatment (for individuals without SCM) in the Current practice comparator was $111,771 to $121,872, the Current practice comparator was dominant (threshold value of 82.6 QALYs) and; when it was $121,872 or greater, Intervention 2 was dominant;

4. **IntV2_NoSCM_NoVP**: When the pre-ventilation cost of treatment (for individuals without SCM) in Intervention 2 was $115,470 to $191,695, Intervention 2 was dominant (threshold value of 81.0 QALYs) and; when it was $191,695 or greater, the Current practice comparator was dominant;

5. **CP_NoSCM_VP**: When the post-ventilation cost of treatment (for individuals without SCM) in the Current practice comparator was $314,238 to $404,774, the Current practice comparator was dominant (threshold value of 82.6 QALYs) and; when it was $404,744 or greater, Intervention 2 was dominant;
6. **IntV1_NoSCM_VP**: When the post-ventilation cost of treatment (for individuals without SCM) in Intervention 1 was $294,805 to $372,019, Intervention 1 was dominant (threshold value of 82.6 QALYs) and; when it was $372,019 or greater, Intervention 2 was dominant;

7. **IntV2_NoSCM_VP**: When the post-ventilation cost of treatment (for individuals without SCM) in Intervention 2 was $309,110 to $404,508, Intervention 2 was dominant (threshold value of 81.0 QALYs) and; when it was $405,508 or greater, the Current practice comparator was dominant;

8. **CP_NoSCM_YV**: When there were 4.61 to 7.02 post-ventilation years of life (for individuals without SCM) in the Current practice comparator, Intervention 2 was dominant (threshold value of 82.6 QALYs) and; when there were 7.02 years or more, the Current practice comparator was dominant;

9. **IntV1_NoSCM_YV**: When there were 7.61 to 10.25 post-ventilation years of life (for individuals without SCM) in Intervention 1, Intervention 2 was dominant (threshold value of 82.6 QALYs) and; when there were 10.25 years or more, Intervention 1 was dominant;

10. **IntV2_NoSCM_YV**: When there were 5.61 to 7.59 post-ventilation years of life (for individuals without SCM) in Intervention 2, the Current practice comparator was dominant (threshold value of 81.0 QALYs) and; for when there were 7.59 years or more, Intervention 2 was dominant.

**Multivariate sensitivity analysis**

**Cost-effectiveness analysis**

A multivariate sensitivity analysis was conducted in Treeage Pro 2009, using the program's multivariate sensitivity analysis function, to examine the impact of the combined uncertainty of the two parameters in the univariate analysis that the results of the CEA were shown to be sensitive to (CP_NoSCM_YV & IntV2_NoSCM_YV). The results of the multivariate analysis are displayed in graphical form in Figure 10 and in tabular form in Table 32.

In the graph in Figure 10, the values of the CP_NoSCM_YV parameter are listed on the x-axis, and the values of IntV2_NoSCM_YV parameter are listed on the y-axis. The graph is partitioned into shaded regions that represent the incremental Cost-effectiveness ratios for the range of possible values for the two parameters and the points at which a threshold of dominance is reached (Briggs et al, 2012). A maximum acceptable ratio of $50,000 (a life year or QALY) was used in the analysis, with the results reported as net health benefits.
The results of the analysis of the combined effects of CP_NoSCM_YV and IntV2_NoSCM_YV were as follows: (i) Intervention 1 was dominant (for a total of 469.61 life-years) when the number of post-ventilation years of life (for individuals without SCM) in the Current practice comparator and Intervention 2 were no greater than their respective minimums of 4.61 & 5.61; (ii) Intervention 2 was dominant (for a total of 473.55 to 582.40 life years) when the number of post-ventilation years of life (for individuals without SCM) in the Current practice comparator was between 4.61 to 7.212 years and between 6.478 and 9.08 years in Intervention 2; and; (ii) the Current practice comparator was dominant (for a total of 496.17 to 604.42 life-years) when the number of post-ventilation years of life (for individuals without SCM) in the Current practice comparator was between 5.4578 to 8.08 years and between 5.61 to 9.08 years in Intervention 2.

In the following table (Table 32) the values of the CP_NoSCM_YV parameter are listed in columns three to seven (from left to right), with the values of the IntV2_NoSCM_YV parameter listed in the leftmost column (column one). The incremental Cost-effectiveness ratios for the combination of the two parameters are listed from the third row down (counting down from the topmost row) with the results for Intervention 1 (IntV2), Intervention 2 (IntV2) and the Current practice comparator (CP) in the fourth, fifth and sixth rows.
(respectively). The dominant intervention or comparator for each combination of the two values is highlighted in bold italics.

**Table 32: Multivariate analysis - CP_NoSCM_YV & IntV2_NoSCM_YV (Life years)**

<table>
<thead>
<tr>
<th>IntV2_NoSCM_YV</th>
<th>CP_NoSCM_YV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.61</td>
</tr>
<tr>
<td>9.08</td>
<td>IntV1</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
</tr>
<tr>
<td></td>
<td>CP</td>
</tr>
<tr>
<td>8.213</td>
<td>IntV1</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
</tr>
<tr>
<td></td>
<td>CP</td>
</tr>
<tr>
<td>7.345</td>
<td>IntV1</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
</tr>
<tr>
<td></td>
<td>CP</td>
</tr>
<tr>
<td>6.478</td>
<td>IntV1</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
</tr>
<tr>
<td></td>
<td>CP</td>
</tr>
<tr>
<td>5.61</td>
<td>IntV1</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
</tr>
<tr>
<td></td>
<td>CP</td>
</tr>
</tbody>
</table>

---

The same format is used for Tables 33 to 38 for the results of the Cost-utility analysis.
Cost-utility analysis

There were four model parameters that were examined in the multivariate sensitivity analysis of the results of the Cost-utility analysis, specifically: (i) post ventilation, utility weighted quality of life (AQoL\_V) and; (ii) the post-ventilation annual cost of treatment (for individuals without SCM) (CP\_NoSCM\_VP, IntV1\_NoSCM\_VP & IntV2\_NoSCM\_VP). These were the four parameters that were shown in the univariate analysis to have the greatest impact on the results of the Cost-utility analysis. A two-way matrix was constructed to map out all of the possible combinations of the four variables (Walkers & Miners, 2005) which were as follows: (i) AQoL\_V & CP\_NoSCM\_VP; (ii) AQoL\_NoV & IntV1\_NoSCM\_VP; (iii) AQoL\_V & IntV2\_NoSCM\_VP; (iv) CP\_NoSCM\_VP & IntV1\_NoSCM\_VP; (v) CP\_NoSCM\_VP & IntV2\_NoSCM\_VP and; (vi) IntV1\_NoSCM\_VP & IntV2\_NoSCM\_VP.

The multivariate sensitivity analysis was undertaken in Treeage Pro 2009 using the method that was previously described for the Cost-effectiveness analysis. The results of the analysis are displayed in graphical form in Figures 11 to 16 and in tabular form in Tables 33 to 38.

The results for the multivariate analysis of the AQoL\_V and CP\_NoSCM\_VP parameters are displayed in graphical form in Figure 11 and in tabular form in Table 33.
The results of the analysis of the combined effects of AQoL_V and CP_NoSCM_VP were as follows: (i) Intervention 1 was dominant (for a total of 111.26 to 143.19 QALYs) when post-ventilation utility weighted HRQoL was 0.2127 to 0.2527 & the post-ventilation cost of treatment (for individuals without SCM) in the Current practice comparator was $314,238 to $523,731; (ii) Intervention 2 was dominant (for a total of 32.78 to 82.62 QALYs) when post-ventilation utility weighted HRQoL was 0.0927 to 0.1727 & the post-ventilation cost of treatment (for individuals without SCM) in the Current practice comparator was $471,358 to $523,731 and; (iii) the Current practice comparator was dominant (for a total of 48.89 to 114.29 QALYs) when post-ventilation utility weighted HRQoL was 0.0927 to 0.2127 & the post-ventilation cost of treatment (for individuals without SCM) in the Current practice comparator was $314,238 to $418,985.
Table 33: Multivariate analysis - AQoL_V & CP_NOSCM_VP (QALYs)

<table>
<thead>
<tr>
<th>CP_NoSCM_VP</th>
<th>AQoL_V</th>
<th>0.0927</th>
<th>0.1327</th>
<th>0.1727</th>
<th>0.2127</th>
<th>0.2527</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ 523,731</td>
<td>IntV1</td>
<td>15.50</td>
<td>47.42</td>
<td>79.34</td>
<td><strong>111.26</strong></td>
<td><strong>143.19</strong></td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td><strong>32.78</strong></td>
<td>57.70</td>
<td><strong>82.62</strong></td>
<td>107.54</td>
<td>132.46</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>26.05</td>
<td>47.85</td>
<td>69.65</td>
<td>91.46</td>
<td>113.26</td>
</tr>
<tr>
<td>$ 471,358</td>
<td>IntV1</td>
<td>15.50</td>
<td>47.42</td>
<td>79.34</td>
<td><strong>111.26</strong></td>
<td><strong>143.19</strong></td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td><strong>32.78</strong></td>
<td>57.70</td>
<td><strong>82.62</strong></td>
<td>107.54</td>
<td>132.46</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>31.76</td>
<td>53.56</td>
<td>75.36</td>
<td>97.16</td>
<td>118.97</td>
</tr>
<tr>
<td>$ 418,985</td>
<td>IntV1</td>
<td>15.50</td>
<td>47.42</td>
<td>79.34</td>
<td><strong>111.26</strong></td>
<td><strong>143.19</strong></td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td><strong>32.78</strong></td>
<td>57.70</td>
<td><strong>82.62</strong></td>
<td>107.54</td>
<td>132.46</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td><strong>37.47</strong></td>
<td><strong>59.27</strong></td>
<td>81.07</td>
<td>102.87</td>
<td>124.68</td>
</tr>
<tr>
<td>$ 366,611</td>
<td>IntV1</td>
<td>15.50</td>
<td>47.42</td>
<td>79.34</td>
<td><strong>111.26</strong></td>
<td><strong>143.19</strong></td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td><strong>32.78</strong></td>
<td>57.70</td>
<td><strong>82.62</strong></td>
<td>107.54</td>
<td>132.46</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td><strong>43.18</strong></td>
<td><strong>64.98</strong></td>
<td><strong>86.78</strong></td>
<td>108.58</td>
<td>130.39</td>
</tr>
<tr>
<td>$ 314,238</td>
<td>IntV1</td>
<td>15.50</td>
<td>47.42</td>
<td>79.34</td>
<td>111.26</td>
<td><strong>143.19</strong></td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td>32.78</td>
<td>57.70</td>
<td>82.62</td>
<td>107.54</td>
<td>132.46</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td><strong>48.89</strong></td>
<td><strong>70.69</strong></td>
<td><strong>92.49</strong></td>
<td><strong>114.29</strong></td>
<td>136.09</td>
</tr>
</tbody>
</table>

The results for the multivariate analysis of the AQoL_V and IntV1_NoSCM_VP parameters are displayed in graphical form in Figure 12 and in tabular form in Table 34.
The results of the analysis of the combined effects of AQoL_V and IntV1_NoSCM_VP were as follows: (i) Intervention 1 was dominant (for a total of 62.71 to 158.47 QALYs) when post-ventilation utility weighted HRQoL was 0.1327 to 0.2527 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 1 was $294,805 to $442,208; (ii) Intervention 2 was dominant (for a total of 82.62 to 132.46 QALYs) when post-ventilation utility weighted HRQoL was 0.1727 to 0.2527 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 1 was $393,074 to $491,342 and; (iii) the Current practice comparator was dominant (for a total of 37.47 to 59.27 QALYs) when post-ventilation utility weighted HRQoL was 0.0927 to 0.1327 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 1 was $294,805 to $491,342.
Table 34: Multivariate analysis - AQoL_V & IntV1_NoSCM_VP (QALYs)

<table>
<thead>
<tr>
<th>IntV1_NoSCM_VP</th>
<th>AQoL_V</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0927</td>
<td>0.1327</td>
<td>0.1727</td>
<td>0.2127</td>
</tr>
<tr>
<td>$ 491,342</td>
<td>IntV1</td>
<td>0.22</td>
<td>32.14</td>
<td>64.06</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td>32.78</td>
<td>57.70</td>
<td>82.62</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>37.47</td>
<td>59.27</td>
<td>81.07</td>
</tr>
<tr>
<td>$ 442,208</td>
<td>IntV1</td>
<td>7.86</td>
<td>39.78</td>
<td>71.70</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td>32.78</td>
<td>57.70</td>
<td>82.62</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>37.47</td>
<td>59.27</td>
<td>81.07</td>
</tr>
<tr>
<td>$ 393,074</td>
<td>IntV1</td>
<td>15.50</td>
<td>47.42</td>
<td>79.34</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td>32.78</td>
<td>57.70</td>
<td>82.62</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>37.47</td>
<td>59.27</td>
<td>81.07</td>
</tr>
<tr>
<td>$ 343,939</td>
<td>IntV1</td>
<td>23.15</td>
<td>55.07</td>
<td>86.99</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td>32.78</td>
<td>57.70</td>
<td>82.62</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>37.47</td>
<td>59.27</td>
<td>81.07</td>
</tr>
<tr>
<td>$ 294,805</td>
<td>IntV1</td>
<td>30.79</td>
<td>62.71</td>
<td>94.63</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td>32.78</td>
<td>57.70</td>
<td>82.62</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>37.47</td>
<td>59.27</td>
<td>81.07</td>
</tr>
</tbody>
</table>

The results for the multivariate analysis of the AQoL_V and CP_NoSCM_VP parameters are displayed in graphical form in Figure 13 and in tabular form in Table 35.
Figure 13: Multivariate analysis - AQoL_V & IntV2_NoSCM_VP (QALYs)

Multivariate sensitivity analysis
AQoL_V & IntV2_NoSCM_VP

The results of the analysis of the combined effects of AQoL_V and IntV2_NoSCM_VP were as follows: (i) Intervention 1 was dominant (for a total of 111.26 to 143.19 QALYs) when post-ventilation utility weighted HRQoL was 0.2127 to 0.2527 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 2 was $412,147 to $515,183; (ii) Intervention 2 was dominant (for a total of 43.25 to 118.00 QALYs) when post-ventilation utility weighted HRQoL was 0.0927 to 0.2127 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 2 was $309,110 to $360,628 and; (iii) the Current practice comparator was dominant (for a total of 37.47 to 81.07 QALYs) when post-ventilation utility weighted HRQoL was 0.0927 to 0.1727 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 2 was $360,628 to $515,183.
Table 35: Multivariate analysis - AQoLV & IntV2_NoSCM_VP (QALYs)

<table>
<thead>
<tr>
<th>IntV2_NoSCM_VP</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$515,183</td>
<td>IntV1</td>
<td>15.50</td>
<td>47.42</td>
<td>79.34</td>
<td>111.26</td>
</tr>
<tr>
<td>$463,665</td>
<td>IntV1</td>
<td>15.50</td>
<td>47.42</td>
<td>79.34</td>
<td>111.26</td>
</tr>
<tr>
<td>$412,147</td>
<td>IntV1</td>
<td>15.50</td>
<td>47.42</td>
<td>79.34</td>
<td>111.26</td>
</tr>
<tr>
<td>$360,628</td>
<td>IntV1</td>
<td>15.50</td>
<td>47.42</td>
<td>79.34</td>
<td>111.26</td>
</tr>
<tr>
<td>$309,110</td>
<td>IntV1</td>
<td>15.50</td>
<td>47.42</td>
<td>79.34</td>
<td>111.26</td>
</tr>
<tr>
<td>CP</td>
<td>37.47</td>
<td>59.27</td>
<td>81.07</td>
<td>102.87</td>
<td>124.68</td>
</tr>
<tr>
<td>CP</td>
<td>37.47</td>
<td>59.27</td>
<td>81.07</td>
<td>102.87</td>
<td>124.68</td>
</tr>
<tr>
<td>CP</td>
<td>37.47</td>
<td>59.27</td>
<td>81.07</td>
<td>102.87</td>
<td>124.68</td>
</tr>
<tr>
<td>CP</td>
<td>37.47</td>
<td>59.27</td>
<td>81.07</td>
<td>102.87</td>
<td>124.68</td>
</tr>
<tr>
<td>CP</td>
<td>37.47</td>
<td>59.27</td>
<td>81.07</td>
<td>102.87</td>
<td>124.68</td>
</tr>
</tbody>
</table>

The results for the multivariate analysis of the CP_NoSCM_VP and IntV1_NoSCM_VP parameters are displayed in graphical form in Figure 14 and in tabular form in Table 36.
The results of the analysis of the combined effects of CP_NoSCM_VP and IntV1_NoSCM_VP were as follows: (i) Intervention 1 was dominant (for a total of 86.99 to 94.63 QALYs) when the post-ventilation cost of treatment (for individuals without SCM) in the Current practice comparator was $314,238 to $523,731 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 1 was $249,805 to $343,939; (ii) Intervention 2 was dominant (for a total of 82.62 QALYs) when the post-ventilation cost of treatment (for individuals without SCM) in the Current practice comparator was $393,074 to $491,342 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 1 was $393,074 to $491,342 and (iii) the Current practice comparator was dominant (for a total of 86.78 to 92.49 QALYs) when the post-ventilation cost of treatment (for individuals without SCM) in the Current practice comparator was $314,238 to $366,611 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 1 was $343,939 to $491,342.
### Table 36: Multivariate analysis - CP_NoSCM_VP & IntV1_NoSCM_VP (QALYs)

<table>
<thead>
<tr>
<th>IntV1_NoSCM_VP</th>
<th>CP_NoSCM_VP</th>
<th>CP_NoSCM_VP</th>
<th>CP_NoSCM_VP</th>
<th>CP_NoSCM_VP</th>
<th>CP_NoSCM_VP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$314,238</td>
<td>$366,611</td>
<td>$418,985</td>
<td>$471,358</td>
<td>$523,731</td>
</tr>
<tr>
<td>$491,342</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IntV1</td>
<td>64.06</td>
<td>64.06</td>
<td>64.06</td>
<td>64.06</td>
<td>64.06</td>
</tr>
<tr>
<td>IntV2</td>
<td>82.62</td>
<td>82.62</td>
<td><strong>82.62</strong></td>
<td><strong>82.62</strong></td>
<td><strong>82.62</strong></td>
</tr>
<tr>
<td>CP</td>
<td><strong>92.49</strong></td>
<td><strong>86.78</strong></td>
<td>81.07</td>
<td>75.36</td>
<td>69.55</td>
</tr>
<tr>
<td>$442,208</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IntV1</td>
<td>71.70</td>
<td>71.70</td>
<td>71.70</td>
<td>71.70</td>
<td>71.70</td>
</tr>
<tr>
<td>IntV2</td>
<td>82.62</td>
<td>82.62</td>
<td><strong>82.62</strong></td>
<td><strong>82.62</strong></td>
<td><strong>82.62</strong></td>
</tr>
<tr>
<td>CP</td>
<td><strong>92.49</strong></td>
<td><strong>86.78</strong></td>
<td>81.07</td>
<td>75.36</td>
<td>69.65</td>
</tr>
<tr>
<td>$393,074</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IntV1</td>
<td>79.34</td>
<td>79.34</td>
<td>79.34</td>
<td>79.34</td>
<td>79.34</td>
</tr>
<tr>
<td>IntV2</td>
<td>82.62</td>
<td>82.62</td>
<td><strong>82.62</strong></td>
<td><strong>82.62</strong></td>
<td><strong>82.62</strong></td>
</tr>
<tr>
<td>CP</td>
<td><strong>92.49</strong></td>
<td><strong>86.78</strong></td>
<td>81.07</td>
<td>75.36</td>
<td>69.65</td>
</tr>
<tr>
<td>$343,939</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IntV1</td>
<td>86.99</td>
<td>86.99</td>
<td>86.99</td>
<td>86.99</td>
<td>86.99</td>
</tr>
<tr>
<td>IntV2</td>
<td>82.62</td>
<td>82.62</td>
<td>82.62</td>
<td>82.62</td>
<td>82.62</td>
</tr>
<tr>
<td>CP</td>
<td><strong>92.49</strong></td>
<td>86.78</td>
<td>81.07</td>
<td>75.36</td>
<td>69.65</td>
</tr>
<tr>
<td>$294,805</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IntV1</td>
<td><strong>94.63</strong></td>
<td><strong>94.63</strong></td>
<td><strong>94.63</strong></td>
<td><strong>94.63</strong></td>
<td><strong>94.63</strong></td>
</tr>
<tr>
<td>IntV2</td>
<td>82.62</td>
<td>82.62</td>
<td>82.62</td>
<td>82.62</td>
<td>82.62</td>
</tr>
<tr>
<td>CP</td>
<td>92.49</td>
<td>86.78</td>
<td>81.07</td>
<td>75.36</td>
<td>69.65</td>
</tr>
</tbody>
</table>

The results for the multivariate analysis of the CP_NoSCM_VP and IntV2_NoSCM_VP parameters are displayed in graphical form in Figure 15 and in tabular form in Table 37.
The results of the analysis of the combined effects of CP_NoSCM_VP and IntV2_NoSCM_VP were as follows: (i) Intervention 1 was dominant (for a total of 79.34 QALYs) when the post-ventilation cost of treatment (for individuals without SCM) in the Current practice comparator was $471,358 to $523,731 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 2 was $463,665 to $515,183; (ii) Intervention 2 was dominant (for a total of 80.24 to 93.08 QALYs) when the post-ventilation cost of treatment (for individuals without SCM) in the Current practice comparator was $314,238 to $523,731 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 2 was $390,110 to $413,147 and; (ii) the Current practice comparator was dominant (for a total of 81.07 to 92.49 QALYs) the post-ventilation cost of treatment (for individuals without SCM) in the Current practice comparator was $314,238 to $418,985 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 2 was $309,110 to $515,183.
Table 37: Multivariate analysis - CP_NoSCM_VP versus IntV2_NoSCM_VP (QALYs)

<table>
<thead>
<tr>
<th>IntV2_NoSCM_VP</th>
<th>CP_NoSCM_VP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$314,238</td>
<td>$366,611</td>
</tr>
<tr>
<td>$418,985</td>
<td>$471,358</td>
</tr>
<tr>
<td>$523,731</td>
<td></td>
</tr>
<tr>
<td>$515,183</td>
<td>$79.34</td>
</tr>
<tr>
<td>$79.34</td>
<td></td>
</tr>
<tr>
<td>$79.34</td>
<td>$81.07</td>
</tr>
<tr>
<td>$75.36</td>
<td>69.65</td>
</tr>
<tr>
<td>$463,665</td>
<td>$79.34</td>
</tr>
<tr>
<td>$79.34</td>
<td></td>
</tr>
<tr>
<td>$79.34</td>
<td>$81.07</td>
</tr>
<tr>
<td>$75.36</td>
<td>69.65</td>
</tr>
<tr>
<td>$412,147</td>
<td>$79.34</td>
</tr>
<tr>
<td>$79.34</td>
<td></td>
</tr>
<tr>
<td>$79.34</td>
<td>$81.07</td>
</tr>
<tr>
<td>$75.36</td>
<td>69.65</td>
</tr>
<tr>
<td>$360,628</td>
<td>$86.66</td>
</tr>
<tr>
<td>$86.66</td>
<td></td>
</tr>
<tr>
<td>$86.66</td>
<td>$81.07</td>
</tr>
<tr>
<td>$75.36</td>
<td>69.65</td>
</tr>
<tr>
<td>$309,110</td>
<td>$93.08</td>
</tr>
<tr>
<td>$93.08</td>
<td></td>
</tr>
<tr>
<td>$93.08</td>
<td>$81.07</td>
</tr>
<tr>
<td>$75.36</td>
<td>69.65</td>
</tr>
</tbody>
</table>

The results for the multivariate analysis of the IntV1_NoSCM_VP and IntV2_NoSCM_VP parameters are displayed in graphical form in Figure 16 and in tabular form in Table 38.
The results of the analysis of the combined effects of IntV1_NoSCM_VP and IntV2_NoSCM_VP were as follows: (i) Intervention 1 was dominant (for a total of 86.99 to 94.63 QALYs) when the post-ventilation cost of treatment (for individuals without SCM) in Intervention 1 was $294,805 to $343,939 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 2 was $309,110 to $515,183; (ii) Intervention 2 was dominant (for a total of 86.66 to 93.08 QALYs) when the post-ventilation cost of treatment (for individuals without SCM) in Intervention 1 was $343,939 to $491,342 & the post-ventilation cost of treatment (for individuals without SCM) in Intervention 2 was $309,110 to $360,628 and; (iii) the Current practice comparator was dominant (for a total of 81.07 QALYs) when the post-ventilation cost of treatment (for individuals without SCM) in Intervention 1 was $343,939 to $491,342 & when the post-ventilation cost of treatment (for individuals without SCM) in Intervention 2 was $309,110 to $515,183.
Table 38: Multivariate analysis - IntV1_NoSCM_VP versus IntV2_NoSCM_VP (QALYs)

<table>
<thead>
<tr>
<th>IntV2_NoSCM_VP</th>
<th>IntV1_NoSCM_VP</th>
<th>$294,805</th>
<th>$343,939</th>
<th>$393,074</th>
<th>$442,208</th>
<th>$491,342</th>
</tr>
</thead>
<tbody>
<tr>
<td>$515,183</td>
<td>IntV1</td>
<td>94.63</td>
<td>86.99</td>
<td>79.34</td>
<td>71.70</td>
<td>64.06</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td>67.41</td>
<td>67.41</td>
<td>67.41</td>
<td>67.41</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
</tr>
<tr>
<td>$463,665</td>
<td>IntV1</td>
<td>94.63</td>
<td>86.99</td>
<td>79.34</td>
<td>71.70</td>
<td>64.06</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td>73.83</td>
<td>73.83</td>
<td>73.83</td>
<td>73.83</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
</tr>
<tr>
<td>$412,147</td>
<td>IntV1</td>
<td>94.63</td>
<td>86.99</td>
<td>79.34</td>
<td>71.70</td>
<td>64.06</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td>80.24</td>
<td>80.24</td>
<td>80.24</td>
<td>80.24</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
</tr>
<tr>
<td>$360,628</td>
<td>IntV1</td>
<td>94.63</td>
<td>86.99</td>
<td>79.34</td>
<td>71.70</td>
<td>64.06</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td>86.66</td>
<td>86.66</td>
<td>86.66</td>
<td>86.66</td>
<td>86.66</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
</tr>
<tr>
<td>$309,110</td>
<td>IntV1</td>
<td>94.63</td>
<td>86.99</td>
<td>79.34</td>
<td>71.70</td>
<td>64.06</td>
</tr>
<tr>
<td></td>
<td>IntV2</td>
<td>93.08</td>
<td>93.08</td>
<td>93.08</td>
<td>93.08</td>
<td>93.08</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
<td>81.07</td>
</tr>
</tbody>
</table>
Probabilistic sensitivity analysis

The following distributions were used in the probabilistic sensitivity analysis: (i) Beta\(^90\), for condition specific HRQoL (MRF28) and generic utility weighted HRQoL (AQoL1); (ii) Uniform\(^91\), for the proportion of deaths due to SCM; (iii) Uniform, for the years spent prior to and after the initiation of LTMV; (iv) Uniform, for the annual cost of treatment and; (v) Uniform, for the discount rate. The mean and standard deviation for pre and post ventilation utility weighted quality of life\(^92\) were used by Treeage Pro 2009 to approximate the alpha and beta values for the AQoL_NoV and AQoL_V parameters. This method could not be used to approximate the values for pre and post ventilation condition specific quality of life (MRF28_NoV & MRF28_V) as it returned a distribution with two modes, one being zero and the other being one. An alpha of 8 and a beta of 1.3 (with an expected value of 0.8602) were instead used for the MRF28_NoV parameter and an alpha of 8 and a beta of 5 (with an expected value of 0.6153) was used for the MRF28_V parameter.

The values that were used to define the uniform distribution in Treeage Pro 2009 were taken from the lower and upper ranges of the model parameters as listed in Tables 27, 28 & 29. A uniform distribution as used, for the previously listed model parameters, for the following reasons: (i) the mean proportion of deaths that were due to severe cardiomyopathy, as the primary cause of death in the study of Eagle et al (2002), were reported with a 95% confidence interval and not a standard deviation as required for the use of a normal distribution; (ii) the years spent prior to and after the initiation of ventilation were also derived from the means and 95% confidence intervals reported in the study of Eagle et al (2002); (iii) the annual cost of treatment was not based on a sampled mean but was derived from an estimate of the annual cost of treatment and a model of service delivery for domiciliary LTMV for a homogenous cohort of 100 individuals with DMD (as reported in Chapter 7).

\(^90\) A binomial distribution (bounded by 0 and 1) that is widely used for proportions (Briggs, 2005; Upton & Cook, 2008).
\(^91\) A distribution in which all of values, within a pre-specified range, are equally likely to occur (Upton & Cook, 2008).
\(^92\) As reported in Chapter 7: (i) AQoL_NoV - mean of 0.1349, standard deviation of 0.0478 and; (ii) AQoL_V - mean of 0.1727, standard deviation of 0.0322).
The probabilistic sensitivity analysis was undertaken through the Monte Carlo Simulation - sampling, Probabilistic sensitivity analysis function in Treeage Pro 2009. The simulation was run for 10,000 cycles with the values of the model parameters for each cycle being randomly sampled from the pre-defined probability distributions.

**Cost-effectiveness analysis**

The results of the probabilistic sensitivity analysis of the results of the Cost-effectiveness analysis are reported in Table 39, as summary statistics of the cost and effectiveness of the Current practice comparator and the two interventions. The results are in presented in order, from the lowest to highest cost.

**Table 39: Probabilistic analysis - Summary of results (Life years)**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Current practice</th>
<th>Intervention 2</th>
<th>Intervention 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>$2,493,622</td>
<td>$2,651,775</td>
<td>$2,956,001</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>$455,095</td>
<td>$485,397</td>
<td>$543,273</td>
</tr>
<tr>
<td>Minimum</td>
<td>$1,369,612</td>
<td>$1,523,971</td>
<td>$1,692,549</td>
</tr>
<tr>
<td>Median</td>
<td>$2,445,950</td>
<td>$2,610,440</td>
<td>$2,916,675</td>
</tr>
<tr>
<td>Maximum</td>
<td>$4,174,260</td>
<td>$4,430,664</td>
<td>$4,929,829</td>
</tr>
<tr>
<td><strong>Effectiveness (Life years)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>555.4</td>
<td>532.5</td>
<td>485.8</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>91.0</td>
<td>94.4</td>
<td>110.3</td>
</tr>
<tr>
<td>Minimum</td>
<td>267.3</td>
<td>239.5</td>
<td>105.3</td>
</tr>
<tr>
<td>Median</td>
<td>550.4</td>
<td>528.3</td>
<td>482.2</td>
</tr>
<tr>
<td>Maximum</td>
<td>920.0</td>
<td>891.1</td>
<td>940.3</td>
</tr>
</tbody>
</table>

The results for Intervention 2 and the Current practice comparator as presented as a scatter plot in Figure 17, with the results for Intervention 1 and the Current practice comparator presented as a scatter plot in Figure 18. In both of the scatter plots a maximum acceptable ratio (of costs to benefits) of $50,000 a life year was used. In the scatter plots the x-axis is the mean difference in effectiveness (for each simulation run) and the y-axis is the mean difference in cost (for each simulation run).
In the scatter plot (Figure 17) of the results of the Probabilistic sensitivity analysis of the CEA, Intervention 2 was; (i) dominated by the Current practice comparator in 28.4% of cases (2824 of the simulations); (ii) more costly and more effective in 26.9% of cases (2690 of the simulations); (iii) less costly and less effective in 0.98% of cases (98 of the simulations) and; (iv) dominant in 4.38% of cases (438 of the simulations).
In the scatter plot (Figure 18) of the results of the Probabilistic sensitivity analysis of the CEA, Intervention 1 was: (i) dominated by the Current practice comparator in 56.5% of cases (5650 of the simulations); (ii) more costly and more effective in 13.28% of cases (1328 of the simulations); (iii) less costly and less effective in 0.14% of cases (14 of the simulations) and; (iv) dominant in 0.73% of cases (73 of the simulations).
The results of the probabilistic sensitivity analysis of the CEA are summarised in a Cost-effectiveness acceptability curve in Figure 19.

**Figure 19: Cost-effectiveness acceptability curve (Life years)**

![CEA: Acceptability Curve](image)

The probability (in Figure 19) that the Current practice comparator will be cost-effective for values up to $100,000 a life-year was approximately 0.50 to 0.74. The probability that Intervention 2 will be cost-effective was approximately 0.28 to 0.33, and for Intervention 1 it was approximately 0.04 to 0.18.
Cost-utility analysis

The results of the probabilistic sensitivity analysis for the CUA are reported in Table 40 as summary statistics of the cost and effectiveness of the Current practice comparator and the two interventions. The results are in presented in order, from the lowest to highest cost.

Table 40: Probabilistic analysis – Summary of results (QALYs)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Current practice</th>
<th>Intervention 2</th>
<th>Intervention 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>$2,499,205</td>
<td>$2,655,313</td>
<td>$2,958,831</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>$463,362</td>
<td>$489,391</td>
<td>$541,866</td>
</tr>
<tr>
<td>Minimum</td>
<td>$1,368,651</td>
<td>$1,500,406</td>
<td>$1,644,940</td>
</tr>
<tr>
<td>Median</td>
<td>$2,454,984</td>
<td>$2,613,617</td>
<td>$2,920,714</td>
</tr>
<tr>
<td>Maximum</td>
<td>$4,239,228</td>
<td>$4,403,237</td>
<td>$4,931,230</td>
</tr>
<tr>
<td><strong>Effectiveness (QALYs)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>124.57</td>
<td>126.34</td>
<td>131.57</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>26.46</td>
<td>26.21</td>
<td>28.43</td>
</tr>
<tr>
<td>Minimum</td>
<td>52.72</td>
<td>53.03</td>
<td>55.26</td>
</tr>
<tr>
<td>Median</td>
<td>122.69</td>
<td>124.26</td>
<td>128.95</td>
</tr>
<tr>
<td>Maximum</td>
<td>251.75</td>
<td>241.59</td>
<td>252.28</td>
</tr>
</tbody>
</table>

The results for Intervention 2 and the Current practice comparator as presented as a scatter plot in Figure 20, with the results for Intervention 1 and the Current practice comparator presented as a scatter plot in Figure 21. In both of the scatter plots a maximum acceptable ratio (of costs to benefits) of $50,000 a QALY was used. In the scatter plots the x-axis is the mean difference in effectiveness (for each simulation run) and the y-axis is the mean difference in cost (for each simulation run).
In the scatter plot (Figure 20) of the results of the Probabilistic sensitivity analysis of the CUA, Intervention 2 was: (i) dominated by the Current practice comparator in 16.25% of cases (1625 of the simulations); (ii) more costly and more effective in 26.79% of cases (2679 of the simulations); (iii) less costly and less effective in 8.52% of cases (14 simulations) and; dominant in 9.69% of cases (969 of the simulations).
Figure 21: Probabilistic analysis - Intervention 1, scatter plot of results (QALYs)

In the scatter plot (Figure 21) of the results of the Probabilistic sensitivity analysis of the CUA, Intervention 1 was: (i) dominated by the current practice comparator in 22.62% of cases (2,262 of the simulations); (ii) more costly and more effective in 33.18% of cases (3318 of the simulations); (iii) less costly and less effective in 4.39% of cases (439 of the simulations) and (iv); dominant in 8.23% of cases (823 of the simulations).
The results of the probabilistic sensitivity analysis of the CUA are summarised in a Cost-effectiveness acceptability curve in Figure 22.

**Figure 22: Cost-effectiveness acceptability curve (QALYs)**

The probability (in Figure 22) that the Current practice comparator will be cost-effective for values up to $100,000 a QALY was approximately 0.35 to 0.52; for Intervention 2 it was 0.29 to 0.33 and; for Intervention 1 is it approximately 0.17 to 0.42.
Summary and discussion

Cost-effectiveness analysis

The results of the Cost-effectiveness analysis were shown in the univariate sensitivity analysis to be sensitive to the number of years of life (post LTMV) in the Current practice comparator and Intervention 2. In the base-case analysis there was a decrease in condition specific quality HRQoL (MRF28) of 28.81% \((n=3)\) after the initiation of ventilation. As previously discussed the early initiation of LTMV (prior to the onset of diurnal hypercapnia) does not provide any additional survival benefit beyond that achieved from its use in Current practice and consequently, all else being equal, the event pathways with the shortest duration of ventilation (i.e. the Current practice comparator & Intervention 2) were found to be those that were the most likely to be cost-effective.

The Current practice comparator was dominant when: (i) there were 4.61 to 6.26 post-ventilation years of life (for individuals without SCM) in the Current practice comparator; (ii) there were 5.61 to 8.33 post-ventilation years of life (for individuals without SCM) in Intervention 2 and; (iii) there were 5.45 to 8.08 post-ventilation years of life (for individuals without SCM) in the Current practice comparator and 5.61 to 9.08 post-ventilation years of life (for individuals without SCM) in Intervention 2. Intervention 2 was dominant when: (i) there were 4.61 to 6.26 years of life in the Current practice comparator (for individuals without SCM) after the initiation of ventilation; (ii) there were 8.33 years of life, or more, in Intervention 2 (for individuals without SCM) after the initiation of LTMV and; (iii) there were 4.61 to 7.212 years of life, post-ventilation for individuals without SCM, in the Current practice comparator and 6.478 to 9.09 years of life, post-ventilation for individuals without SCM, in Intervention 2. Intervention 1 was dominant when there were no more than 4.61 and 5.61 years of life post-ventilation years for individuals, without SCM, in the Current practice comparator and Intervention 2 respectively.
Cost-utility analysis

In the base-case of the Cost-Utility analysis, Intervention 1 and 2 produced only a small increase in the number of QALYs, above that of the comparator, of 5.54 and 2.67 QALYs and there were no strategies that were clearly dominated by any other, in either standard or extended dominance. A result that was shown to be sensitive to; (ii) the value of pre and post ventilation generic utility weighted HRQoL (AQoL); (ii) the pre-ventilation cost of treatment; (iii) the post-ventilation cost of treatment; and (iv) the number of years of life after the initiation of ventilation. There was 21.89% (n=3) improvement in generic utility weighted HRQoL (AQoL) after the initiation of LTMV and this is reflected in the incremental QALY gain seen in Intervention 1 and 2.

In the sensitivity analysis the Current practice comparator was dominant when: (i) pre-ventilation HRQoL was no more than 11.19% below its base case value; (ii) there was no more than a 10.07% increase in post-ventilation HRQoL; (iii) there was an improvement in post-ventilation HRQoL of no more than 36.58% and the cost of treatment (post LTMV), in the Current practice comparator, was no greater than its base-case value; (iv) there was a decline in HRQoL following the initiation of ventilation of at least 1.66% and the cost of treatment (post LTMV) in Intervention 1 was no more than 25.00% below, and no more than 20.00% above, its base-case value; (v) there was an improvement in post-ventilation HRQoL of no more than 21.89% and the cost of treatment (post LTMV) in Intervention 2, was between no more than 8.25% below and no more than 23.70% above its base-case value.

Intervention 2 was dominant when: (i) pre-ventilation HRQoL was 10.07% to 12.42% below its base-case value; (ii) there was an increase of post-ventilation HRQoL of between 10.07% to 29.00%; (iii) there was no more than a 21.89% increase in post-ventilation HRQoL and the cost of treatment (post LTMV) in the Current practice comparator was at least 11.11% above its base-case value; (iv) there was at least a 21.89% increase in post-ventilation HRQoL and the cost of treatment (post LTMV) in Intervention 1 was not less than its base-case value and; (ii) there was an increase of post-ventilation HRQoL of no more than 36.58% and the cost of treatment (post LTVM) in Intervention 2 was 8.25% to 21.36% below its base-case value.

Intervention 1 was dominant when: (i) pre-ventilation HRQoL was 11.05% to 25.87% below its base-case value; (ii) there was an increase of post-ventilation HRQoL of at least 29.00%; (ii) there was a post-ventilation increase in HRQoL of at least 36.58%, and the cost of treatment (post LTMV) in the Current practice comparator was no less than 25% below, and no more than 20% above, its base-case value; (iii) there was a post-ventilation reduction in HRQoL of no more than 1.66% and the cost of treatment (post
LTMV) in Intervention 1 was not more than 27.54% below, and not more than 10.88% above its base-case value and; (iv) there was a post-ventilation increase in HRQoL (AQoL) of at least 36.58% and the cost of treatment (post LTMV) in Intervention 2 was 4.63% to 23.70% above its base-case value.
Conclusion

In the base-case of the Cost-effectiveness analysis the initiation of LMTV from the onset of SBD in REM without hypercapnia (Intervention 1) and the initiation of LTMV from the onset of SBD in REM with hypercapnia (Intervention 2) were dominated by the Current practice comparator. This was a result of their being a decrease in condition specific quality HRQoL (MRF28) of 28.81% \((n=3)\) after the initiation of LTMV. The probability that the Current practice comparator would be cost-effective for values up to $100,000 a life-year was approximately 0.50 to 0.74; for Intervention 2 it was approximately 0.28 to 0.33, and for Intervention 1 it was approximately 0.04 to 0.18.

There were no strategies in the Cost-utility analysis that were clearly dominated by any other, in either standard or extended dominance. There was 21.89\% \((n=3)\) improvement in generic utility weighted HRQoL (AQoL) after the initiation of LTMV and this is reflected in the incremental QALY gains seen in Intervention 1 and 2. The probability that the Current practice comparator would be cost-effective for values up to $100,000 a QALY was approximately 0.35 to 0.52; for Intervention 2 it was 0.29 to 0.33 and; for Intervention 1 is it approximately 0.17 to 0.42.

In the next and final chapter of the thesis there is a summary of the results of the economic evaluation, its limitations and how the results compare to those of related economic evaluations. The distributive implications are then discussed, along with the need for further research.
Chapter 9: Discussion

Introduction

The previous chapters of the thesis the reader was provided with an introduction to DMD, its natural history and management, the methods used in economic evaluations, the evidence for the effectiveness of the use of LTMV to support respiratory function, the staffing and equipment requirements of a domiciliary LTMV service and the economic evaluation and its results. The role of this chapter in the thesis is to summarize the results of the evaluation and to provide the reader with a guide to their interpretation.

To achieve this aim the chapter starts with a summary of the results of the economic evaluation. This is followed by the limitations of the evaluation and the generaliability and transferability of the results. There is a comparison of the results of the evaluation to the findings of studies of the use of LTMV in Amyotrophic lateral sclerosis (Motor neuron disease). This is followed by an examination of the distributive implications of economic evaluations for interventions for individuals with rare diseases, the trade-off between efficiency and equity, and the potential for double jeopardy in the strict application of the QALY framework. The public preference for interventions that can avert or delay the death of seriously ill individuals is then examined and the development, in Australia, of public policy frameworks that take this into account. This is followed by an examination of how the difficulties in obtaining a statistically significant sample in studies conducted in rare disease populations could be addressed through the use of innovative research designs and bio statistical techniques, and the use of disease registries.

The chapter ends with a discussion of how the many limitations of the evaluation could be addressed through further research, such as: (i) a replication and extension of the systematic literature review, conducted by two researchers, that included full professional translations of foreign language studies and; (ii) an Australia-wide, multi-institution longitudinal prospective study of the HRQoL of individuals with a DNA confirmed diagnosis of DMD.
Summary of the results

In the base-case of the Cost-effectiveness analysis (all costs and outcomes discounted by 3%) the initiation of LTMV from the onset of SDB in REM without hypercapnia (Intervention 1) and the initiation of LTMV from the onset of SDB in REM with hypercapnia (Intervention 2) were dominated by the Current practice comparator. This was a due to: (i) the early initiation of LTMV in Intervention 1 and 2 providing no additional survival benefit above the use of the intervention in the comparator; (ii) the use of LTMV by individuals with severe cardiomyopathy not being associated with any prolongation of survival; (iii) the decrease in mean post-ventilation condition specific HRQoL (MRF28) of 28.81% (n=3) and; (v) the increase in the duration of LTMV in Intervention 1 and 2 and the associated costs. As all else being equal, given the impact of the intervention on condition specific quality of life, the treatment pathway that had the shortest duration of ventilation, in this case the Current practice comparator, was the one that was most likely to be the most effective, at the least cost.

This conclusion is supported by the univariate and multivariate sensitivity analysis, where the results of the analysis were shown to be sensitive to the duration of ventilation (for individuals without SCM) in the Current practice comparator and Intervention 2. It should be noted that the duration of ventilation in Intervention 2 (in the base-case analysis) is only one year greater than that of the comparator and that there were no individuals with severe cardiomyopathy in Intervention 2 who are ventilated prior to their death, which is also the case in the comparator. In the probabilistic sensitivity analysis, for a maximum acceptable ratio of $100,000 a life-year, the likelihood that the Current practice comparator would be cost-effective was approximately 0.50 to 0.74; for Intervention 2 it was approximately 0.28 to 0.33 and; for Intervention 1 it was 0.04 to 0.18.

In the base-case of the Cost-Utility analysis (all costs and outcomes discounted by 3%), in which there was an increase of mean post-ventilation utility weighted HRQoL (AQoL) of 21.89%, Intervention 1 and 2 produced small incremental increases in the total number of QALYs, of 5.54 and 2.67 respectively. There were no strategies that were clearly dominated by any other, in either standard or extended dominance. The results of the analysis were shown, in univariate and multivariate sensitivity analysis to be sensitive to; (i) the value of pre and post-ventilation generic utility weighted HRQoL (AQoL); (ii) the cost of treatment, prior to and after the initiation of ventilation and; (iii) the duration of ventilation. In the probabilistic sensitivity analysis, for a maximum acceptable ratio of $100,000 a life-year, the likelihood that the Current practice comparator would be cost-effective was approximately 0.35 to 0.52; for Intervention 2 it was approximately 0.29 to 0.33 and; for Intervention 1 it was approximately 0.17 to 0.42.
As the early initiation of LTMV, prior to the onset of diurnal hypercapnia provides no additional survival benefit above its use in current practice, for Intervention 1 and 2 to be cost-effective there needs to substantial increase in post-ventilation HRQoL to offset the costs associated with the increase in the duration of ventilation and the number of ventilated individuals.

As previously stated, the limited evidence for the effectiveness of the intervention, the inability of the small scale study to recruit and retain a statistically significant sample and the use of a researcher developed model of LTMV service delivery, mean that the results of the economic evaluation cannot be seen to be anything other than exploratory in nature.

Study limitations

The decline of respiratory function, by age, that was used in the model for the timing of the initiation of the intervention was obtained from a single study, that of Hahn et al (1997). At the time the decision analytic model was constructed the study of Hahn et al (1997) was the best available and most detailed description of the decline in respiratory function for a cohort of individuals with DMD. As previously described, respiratory function in DMD reaches a plateau after the loss of ambulation from where it declines until the death of the individual, from either cardiac or respiratory failure (as a primary causes of death). In the cohort in the study of Hahn et al (1997) respiratory function reached this plateau between 13-14 years of age, when mean VC was 2658 mls (SD: 870) and within a year this had declined by 30% (808 mls) to a mean of 1836 mls (SD: 994). While the same pattern of respiratory function reaching a plateau, and then declining is reported in other studies of respiratory function in DMD, such as that of Lissoni (2001), the mean max VC reported in this studies varies as a function of the age at which ambulation is lost. For example, in the study of Lissoni et al (2001) respiratory function reached this plateau at approximately 12 years of age, when mean VC was only 1850 mls (SD: 850). The variation in the mean age at which ambulation was lost in these two studies could be the result of changes made over time in the criteria used to establish a diagnosis of DMD, which may have resulted in differences in the number of individuals in the two studies with other forms of muscular dystrophy that had been that had been misclassified as DMD (i.e. Becker muscular dystrophy or Limb Girdle muscular dystrophy Type 21). It is also possible that there were differences in the health status of the two cohorts that may have impacted on their respiratory function, as the study of Lissoni (2001) was conducted in an inpatient population and the study of Hahn et al (1997) was conducted in an outpatient population. It is not known to what degree the decline of respiratory function reported in Hahn et al (1997), or that reported in Lissoni (2001) is representative of the decline of respiratory function by age in an Australian cohort of individuals with DMD.
The event pathways for the current practice management of DMD in the economic evaluation were taken from best practice guidelines that were developed by clinicians from the United Kingdom and United States. There is very little available information about the current practice management of DMD in Australia and consequently it is difficult to determine the extent to which these guidelines reflect the management of the disease in Australia. The more that the event pathway deviate from the current practice management of DMD in Australia, the less reliable are the estimates of the cost of treatment that were used in the economic evaluation. As previously discussed, the lack of an existing disease specific estimate of the cost of providing domiciliary LTMV in Australia meant that a model of service delivery had to be constructed from government reports and published literature that could be used to generate a cost estimate for use in the economic evaluation. While this was based on the service delivery model used by the Victorian Respiratory Support Service (VRSS) it is not known to what extent it is a reasonable approximation of the cost of providing LTMV to individuals with DMD in Australia.

Economic evaluations are, in most cases, conducted within specific decision-making contexts and an intervention that is known to be cost-effective in Australia will not necessarily be cost-effective in another country. This is not simply due to differences in the prices of goods and services between countries but differences in clinical practice and the alternative treatments that are available for use. The effectiveness of an intervention can vary between countries and settings, for while the clinical effect of a drug on a patient may be relatively unchanged, the same cannot be said for surgical procedures performed by two surgeons in two different countries (Drummond et al, 2005; Roberts et al, 2012).

The use of LTMV in DMD is ultimately based on long-standing clinical experience and expert consensus, as opposed to clear evidence of effectiveness from a series of well conducted randomised controlled trials. The clinical experience on which this consensus is based was gained from not only the use of LTMV in DMD but its use in general for a range of neuromuscular diseases and chest wall disorders. Furthermore, it is based on generalisations made from the use of LTMV in a range of different settings (i.e. inpatient, outpatient) delivered through the use of different forms of mechanical ventilation (i.e. non-invasive negative pressure, invasive positive pressure & non-invasive positive pressure ventilation) and the use of a range of different interfaces and machines (i.e. nasal masks, mouthpiece seals, volume cycled ventilators and pressure cycled ventilators). The findings of this economic evaluation should not be generalised to the use of domiciliary non-invasive positive pressure LTMV, or other forms of LTMV, for populations of individuals with neuromuscular or chest wall disorders, as: (i) the effectiveness of evidence for the use of LTMV is DMD specific and relates only to the use of non-invasive positive pressure ventilation delivered via a nasal mask; (ii) the decision analytic model was constructed around the natural
history of the disease, and the decline of respiratory function by age in DMD and; (iii) the event pathways for the clinical management of the disease are DMD specific.

The results of related economic evaluations

The only economic evaluations of the use of non-invasive LTMV for individuals with neuromuscular diseases or chest wall diseases are for its use in Amyotrophic lateral sclerosis (motor neuron disease). Amyotrophic lateral sclerosis (ALS) is an adult onset (40 to 60 years of age) rapidly progressive invariably fatal disease in which the most common cause of death is respiratory failure. In the majority of cases it is not until there is severe respiratory muscle weakness and breathlessness that the diagnosis is made, at which point the individual has less than a year to live (Gruis, Chernew & Brown, 2005).

Gruis, Chernew & Brown (2005) conducted a Cost-utility analysis, from a societal perspective, of the use of LTMV to support respiratory function in individuals with ALS. The aim of the evaluation was to determine the improvement in HRQoL that would be required for the initiation of LTMV at the time of diagnosis to be cost-effective, as compared to its initiation in current practice at the point when VC falls to less than 50% of predicted (for age & sex). At the time the evaluation was conducted the optimal timing of the initiation of LTMV in ALS was unknown but from the available evidence it did appear that its use prolonged survival and improved quality of life. A Markov process model was used in the evaluation, in which the decline of respiratory function in a hypothetical cohort of individuals with ALS was represented in five health states: (i) mild respiratory insufficiency; (ii) moderate respiratory insufficiency; (iii) severe respiratory insufficiency; (iv) terminal respiratory insufficiency and; (v) death. The only allowable transitions between these health states were from a less severe form of respiratory function, to a more severe form (i.e. from mild respiratory insufficiency to moderate respiratory insufficiency) with death as an absorbing state. The members of the cohort progress through these five health-states as their respiratory function declines. The probabilities for the transitions between the health-states were obtained from the published literature. The authors failed to clearly document the length of the Markov cycle over which these transitions were made.

The primary outcome of the model was the number of QALYs that were produced by the intervention and the comparator. The utility values for each of the health states was obtained from a study conducted by Kiebert et al (2001) that used the EuroQoL EQ-5D to measure HRQoL in 77 individuals with ALS. The only items of resource use that were considered in the evaluation were the cost of the equipment and consumables needed to provide non-invasive positive pressure ventilation. The unit costs for resource usage was obtained from US sources and reported in 2004 US dollars. The time horizon for the
evaluation was 1 year, the mean post-diagnosis lifespan of an individual with ALS. There was no
discounting of costs and outcomes, as this was not necessary given the time horizon of the study. The
incremental cost-utility of the early initiation of LTMV in ALS was found to be $33,800 per QALY ($
62,359, AU 2012\textsuperscript{93}). In a one-way uncertainty and sensitivity analysis the cost-utility of the intervention
remained below $50,000 per QALY ($92,247, AU 2012) across the plausible range for the transition
probabilities, quality of life scores and costs. The authors concluded that the early initiation of LTMV in
ALS could be cost-effective if its use was associated with an improvement in utility weighted quality of life
of 7% to 14% (Gruis, Chernew & Brown, 2005).

The National Institute of Clinical Excellence (NICE) conducted two modelled Cost-effectiveness analyses
of the use of non-invasive pressure LTMV in the management of respiratory insufficiency in ALS. In these
evaluations the use of LTMV, as a component of the management of ALS, was compared to the current
practice management of the disease which at that time did not include the use of LTMV. As in the study of
Gruis, Chernew and Brown (2005) the utility weights for the health-states in the two models were taken
from the study of Kiebert et al (2001). For the purposes of the evaluation it was assumed that LTMV was
used for at least 4 hours every day. The evaluations were undertaken from a health sector perspective
with all costs expressed in 2007-2008 UK pounds. The discount rate that was used in the evaluations and
whether this was applied equally to both costs and outcomes was not reported (National Institute for
Health and Clinical Excellence, 2010).

The following items of resource use were considered in the two models: (i) ventilators; (ii) training in the
use of the ventilators; (iii) ventilator maintenance; (iv) consumables such as masks and headsets and; (v)
batteries and chargers for use in emergencies (or portable use of the ventilators). In the estimation of the
cost of resource use for the intervention ventilator users were supplied with two sets of tubing, two masks

\textsuperscript{93} Costs were converted to 2012 AU dollars using a tool developed by Shemilt, Thomas and Morciano (2010) which
was accessed through the following website http://eppi.ioe.ac.uk/costconversion/default.aspx last accessed
2/9/2013. For a meaningful comparison to be made between the costs of different programs conducted at different
times and in different countries, costs must be converted to a common currency and reference year. The tool does
this in two stages: (i) it converts the cost estimate from its original reference year to the target reference year (i.e.
2012) using a gross domestic product deflator index source from International Monetary Fund and; (ii) it does a
currency conversion using purchasing power parity to adjust for differences in pricing levels between countries, that
uses the prices of a common set of goods and services (Shemilt, Thomas & Morciano, 2010).
and two sets of headgear per annum. Based on expert opinion the estimated working life of a ventilator was five years and 2 years for batteries and chargers. The ventilators, batteries and chargers were costed on an annual equivalent basis using a discount rate of 3.5% (National Institute for Health and Clinical Excellence, 2010).

The mean cost of the initial training in the use of the ventilator and mask was estimated to be 154 pounds with a range of 76 to 232 pounds ($403, AU 2012 range of $199 to 607) based on an mean training period of 3.8 hours per ventilator user and a wage rate of 20 to 61 pounds ($52 to 160, AU 2012) per hour dependent on the qualifications of the health professional. The provision of ventilators and consumables was estimated to cost 543 pounds ($1421, AU 2012) for each two month cycle of the model and 3149 pounds per annum ($8,242, AU 2012). There was no allowance made in the cost-estimates for the re-use of ventilators and other equipment. The cost of the current practice comparator was difficult to determine as there was no consensus in the NICE expert advisory group on what constitutes standard care for individuals with ALS in the United Kingdom and this was ultimately obtained from a study conducted by Munsat (1998) (National Institute for Health and Clinical Excellence, 2010).

A Markov chain model with a three year time horizon was developed for use in the first of the two evaluations. The structure of the model was informed by expert opinion and the approach used in previous evaluations, such as that of Gruis, Chernew & Bush (2005). There were four health states in the model that were used to represent the decline of respiratory function in the disease: (i) mild; (ii) moderate; (iii) severe; and (iv) terminal, with death as the fifth and absorbing state. The probabilities for the transitions between these health states were obtained from Stewart (2000). The cycle length of the model was two months, with the only possible transitions between health states being from a less severe health state, to a more severe health state. In the base-case analysis there was an increase of 0.28 years in expected survival in the group treated with LTMV and an associated QALY gain of 0.201. The cost per QALY for a time horizon of two years was 25,350 pounds ($66,353, AU 2012) and for five years it was 20,666 pounds per QALY ($54,092, AU 2012). A probabilistic sensitivity analysis was conducted in which it was found that for a threshold value of 30,000 pounds per QALY ($78,524, AU 2012) there was a 90% probability that the intervention was cost effective. In the second of the two evaluations a simple decision-analytic model was developed to mirror the results of an RCT conducted by Bourke et al (2006) that investigated the impact that LTMV had on survival and quality of life. There were 41 individuals in the study with respiratory failure (secondary to ALS) who were randomized to receive either standard care, or standard care and LTMV. In this evaluation the use of LTMV was estimated to cost between $13,237 to 30,438 pounds a QALY ($34,648 to 79,671 AU 2012) (National Institute for Health and Clinical Excellence, 2010).
There are important differences in the above listed studies that need to be taken into account in any comparison to the findings of the economic evaluation in the thesis. The time horizon in the studies was only 1 to 3 years due to the relatively short period of time between diagnosis and death in ALS. In the two evaluations conducted by the National Institute for Health and Clinical Excellence the use of LTMV was compared to the current practice management of the disorder in which there was no use of any form of mechanical ventilation. The only evaluation that examined the costs and outcomes associated with the timing of the initiation of LTMV was that of Gruis, Chernew and Brown (2005). While this study was reported to have been conducted from a societal perspective, the only costs and outcomes that were considered were the equipment and consumables required to provide LTMV and its impact on health related quality of life. Notwithstanding, the incremental cost per QALY of between $34,768 to $79,671 (AU, 2012) reported in the three evaluations compares favourably to the incremental cost per QALY for the use of LTMV in DMD of $20,968 for Intervention 2 and $79,548 for Intervention 1 (2012 AU$, discount rate of 3%).

**Distributive implications**

The results of Cost-effectiveness and Cost-utility analyses are commonly expressed in terms of the cost per unit of outcome without any consideration of the characteristics, or circumstances, of the individuals who incur the costs or receive the benefits of an intervention. From a policy making perspective this is seen as a major shortcoming of economic evaluations as there are many situations where the efficient use of resources is seen to be less important than the need for distributinal equity. This is a concern, within clinical decision making, health care policy and normative economics for the fair and impartial distribution of goods and services. Normative economics is an area of economic theory that focuses on how economies should be run. Its prime focus is on the trade-offs that need to be made between efficiency and equity to be able to achieve socially desirable goals. It starts by asking whether current goals could be met through the use of fewer resources (through an increase in efficiency) with the aim of releasing resources that can be used to achieve other desirable objectives. It then considers the equity impacts of these changes and whether the resultant distributions of costs and benefits are acceptable in relation to value judgements about issues such as: the fairness of income and the distribution of wealth; opportunity and; the need for and access to healthcare. A distribution that treats all of the members of a society as equal, for example in relation to their access to healthcare, is one that is said to be horizontally equitable. A distribution that treats people with different needs differently is said to be vertically equitable. In this later distribution, the individuals with the greatest need for access to healthcare receive the most care. The extent to which their needs are met is dependent on the value judgements of decision makers and society.
in general (Siegel, Weinstein & Torrance, 1996; Hurley, 2000; Culyer, 2005; Drummond et al, 2005; Black, Hashimzade & Myles, 2009).

In recent years there have been doubts expressed about the applicability of the standard methods of economic evaluation for interventions for individuals with rare diseases. These doubts relate to the need for clear evidence of effectiveness (as discussed in Chapter 5) and the methods that are used for the valuation of health gain. In Cost-utility analysis there is an underlying assumption that all life-years and QALYs are of equal value, regardless of the characteristics or circumstances of the individuals who receive them. While this approach may appear to be egalitarian and ethically justifiable, it has the potential to seriously disadvantage individuals with chronic diseases and permanent disabilities. It does this by placing them in a position of double jeopardy, as not only are they misfortunate enough to suffer from an illness or disability but because of their potential to benefit their access to interventions that can improve their quality of life, or extend their life expectancy is reduced. For if all else is equal, interventions that save or prolong the lives of individuals who can be returned to full health and normal life expectancy have the potential to produce a greater QALY gain than can interventions for individuals with chronic illnesses or permanent disabilities, for who a return to full health and normal life expectancy is not possible (Harris, 1987; Singer et al, 1995; Garber et al, 1996; Gold et al 1996; Culyer, 2005; Drummond et al, 2005; Drummond et al, 2009; Laupacis, 2009; McCabe, Edlin & Round, 2010; Panju & Bell, 2010).

In Australia, government funded services such as the Pharmaceutical Benefits Scheme (PBS) have begun to develop policy frameworks that explicitly take into account the rarity of a disease or condition, its severity and the availability and effectiveness of alternative treatments. These frameworks explicitly acknowledge the value that Australian society places on the need for vertical equity in the healthcare system and the importance of the rule of rescue. In situations where the total budget impact is expected to be small there is an explicit trade-off between the need for efficiency and a desire to reflect societal values, with more value placed on the health gain produced by an intervention for a rare disease than that for a more common disease. Unfortunately this is not a simple or straightforward process when the effectiveness of an intervention is surrounded by significant uncertainty, the benefit it produces is small, there is no valid alternative and the cost and incremental Cost-effectiveness is high. While these frameworks are a pragmatic policy response to avoid a blanket ban on interventions that are unable to meet the standard requirements for government funding (i.e. less than $ 50,000 AU a QALY) there is uncertainty about the long-term sustainability of these policies due to the budgetary impact of an increasing number of interventions that meet the eligibility criteria (Hauck, Smith & Goddard, 2004; McKie & Richardson, 2005b; Drummond et al, 2009; Laupacis, 2009; McCabe, Edlin & Round, 2010; Panju & Bell, 2010).
The need for further research

There are significant gaps in what is known about many rare diseases, which is at least partly a consequence of the challenges encountered in the recruitment and retention of study participants and the associated difficulty, within a single institution or country, of obtaining a large enough sample for the results of a RCT to be able to reach statistical significance. A difficulty that has led, in recent years, to the development of innovative research designs and biostatical techniques that maximise the data that can be obtained from studies for which there are only a small number of potential study participants. The innovative approaches to research designs for these diseases have evolved from a need to look beyond classical research methodology and includes the use of methods such as: (i) adaptive (Bayesian\textsuperscript{94}) group randomisation; (ii) longitudinal data collection and repeated measurement designs; (iii) group sequential and adaptive research designs\textsuperscript{95} and; (iv) meta-analysis. The innovative biostatical techniques, which are used to increase the efficiency of standard statistical analysis include: (i) Bayesian statistics; (ii) ranking and selection; (iii) non-randomised risk-based group allocation and; (iv) statistical prediction. It is important that no matter what type of research design, or form of statistical inference is used in studies conducted in rare disease populations that it is acceptable to decision-makers and is able to meet the same standards of rigour that are required in studies for more commonly occurring diseases, such as Type 2 diabetes mellitus (Griggs et al, 2009; McCabe, Edlin & Round, 2010; Gerb & Kopcke, 2010; Panju & Bell, 2010; Gliklich & Leavy, 2011).

The data collected in studies conducted in rare disease populations may also be supplemented with information obtained through the use of other methods, such as disease registries. A disease registry is a detailed database of individuals with a given disease that is used to monitor their health status and to provide a resource that can be used to identify potential study participants for a clinical trial. Disease registries also provide a means through which uniform data can be collected to map out the natural history of a disease, the effect of an intervention and the outcome it produces, and its long-term Cost-

\textsuperscript{94} An approach to statistical inference that is based on Baye's theorem in which hypotheses are formed on the basis of prior knowledge about the likelihood of an event. In the standard “frequentist” theory of statistical inference probabilities can only be attached to repeatable events. In Bayesian statistics, probability is used to describe the uncertainty surrounding what is known about the likelihood of a future event based on prior knowledge of its occurrence. This uncertainty may be due to either imperfect knowledge about the event, or its random/unpredictable nature). The test of a hypothesis in Bayesian statistics is an assessment of the probability of which of the null or the alternative hypothesis is more likely to be correct, based on prior knowledge (O’Hagan & Luce, 2003; Black, Hashimaze & Myles, 2009).

\textsuperscript{95} A research design in which study data is monitored as it is being collected, with the results of a series of interim analyses used to: (i) modify the sample size requirements of the study or; (ii) to end a study, once it is clear that the intervention under examination is either effective or ineffective. To keep the risk of a type 1 error within acceptable limits the study methods must be determined in advance, and strictly adhered to as must the conditions that must be met for a intervention to be seen to be effective (Kelly et al, 2005).
effectiveness. The usefulness of disease registries is limited by the potential for bias if there are changes over time in the diagnostic criteria used to classify the disease of interest (i.e., such as a move to the use of DNA analysis to confirm the diagnosis of the disease) or in the natural history and/or severity of the disease as a result of changes in clinical practice (i.e., such as the use of aggressive chest physiotherapy and antibiotics to treat chest infection in DMD). Furthermore, to be able to undertake an economic evaluation for an intervention for a rare disease, data must be collected for both the costs and outcomes of intervention and those of a valid comparator. While there are many disease registries that are established to collect information about the costs and outcomes for a specific intervention, such as a drug, the same degree of effort is not necessarily made to collect the comparative data for an alternative treatment that is required in a Cost-effectiveness analysis. These limitations can be at least partially addressed through the careful design of these registries, the use of study designs that reduce the effects of bias and confounding and statistical techniques to account for their possible effect on study outcomes (Panju & Bell, 2010; Gliklich & Leavy, 2011; Simoons, 2011).

The many limitations of this economic evaluation could be addressed through further research, such as: (i) a study of the current management of DMD in clinical practice in Australia and how this compares to best practice guidelines; (ii) a replication and extension of the systematic literature review, conducted by two researchers, that included full professional translations of foreign language studies; (ii) the establishment of a well-designed Australia wide disease registry for individuals with a DNA confirmed diagnosis of DMD that has as one of its aims, a prospective long-term study of the decline of respiratory function; (v) a Australia-wide, multi-institution longitudinal prospective study of the HRQoL of individuals with a DNA confirmed diagnosis of DMD and; (vi) an Australia-wide, multi-institution analysis of the cost of providing domiciliary LTMV to individuals with DMD. Although respiratory failure is the primary cause of death in DMD, in recent years there have been a number of reports (i.e., Ballard et al., 2012) in which there is an increase in the number of deaths due to cardiac causes in individuals whose respiratory function is supported by LTMV. This would need to be taken into account in any future economic evaluations of the use of the intervention.

As previously discussed it appears that the MRF28 may not have been well accepted by the study participants. This may have been due to the narrow focus of the instrument on the disabling effects of respiratory muscle weakness, in a group who have severe generalized muscle weakness that regardless of their level of respiratory function has left them unable to undertake many of the activities listed in the questionnaire. It is also possible that many individuals found the MRF28 difficult to complete, for while the decline of respiratory function and the need for LTMV is known to be inevitable in DMD, it is not
necessarily an aspect of the disorder that they wish to be reminded of. This may also be an alternative explanation for why the AQoL1 appeared to be better accepted by the study participants than the MRF28.

In retrospect, the impact that declining respiratory function has on condition specific quality of life may have better addressed through the use of an indirect measure, such as a questionnaire for the assessment of sleep quality and duration. The use of such a questionnaire may also provide a mechanism through which it may be possible to separate out the impact that sleep disordered breathing has on quality of life, from other influences, such as the presence of severe generalised muscle weakness. The study would need to be conducted across a number of centres in Australia for there to be any possibility of it being able to detect a statistically significance difference in the HRQoL of individuals with DMD prior to, and after, the initiation of LTMV. Ideally, the study would be conducted using a longitudinal design, with HRQoL measured at six monthly intervals from the loss of ambulation. By doing this it would be possible to map changes in HRQoL of life against the decline in respiratory and cardiac function and other events in the lives of the individuals. To be able to do this would require the data be collected in a manner that minimises the burden that is placed on the individuals, while maximising the recruitment and retention of participants. One possible strategy would be to undertake to collect this data at the six monthly multidisciplinary team reviews and in the later stages of the disease (i.e. post the initiation of LTMV) it could be collected by the outreach nurses as part of their six monthly assessment of the general health of ventilator users.
Conclusion

The aim of the thesis was to produce an estimate of the Cost-effectiveness and Cost-utility of the use of domiciliary long-term non-invasive positive pressure ventilation to support respiratory function in individuals with Duchenne muscular dystrophy. In the economic evaluation the costs and outcomes of two alternative treatment pathways for the timing of the initiation of LTMV in DMD were examined: (i) from the onset of SDB in REM, without hypercapnia at 15.5 years (Intervention 1) and; (ii) from the onset of SDB in REM, with hypercapnia at 17.5 years (Intervention 2) as compared to its initiation in current practice at 18.5 years of age.

To achieve this aim: (i) a formal systematic literature was undertaken to find the best available evidence for the effectiveness of the intervention; (ii) a service delivery model was developed for the provision of domiciliary LTMV to provide a basis for a cost estimate that could be used in the economic evaluation; (iii) a small scale study was conducted to attempt to measure the condition specific and generic utility weighted health related quality of life (HRQoL) of individuals with DMD, prior to and after the initiation of LTMV and; (iv) a decision analytic model was developed to combine the evidence of effectiveness for the intervention, with the cost estimates for service provision and the HRQoL data.

A total of 57 papers were examined in detail in the formal systematic literature review, with only 4 studies meeting all of the inclusion criteria. There were no studies that examined quality of life as an outcome measure. There was only one study that reported the mean life-expectancy of a group treated with LTMV, and a group who were not. As the quantitative pooling of the findings of studies that use non-randomized research designs is not recommended, the evidence for the effectiveness of intervention was obtained from a single study that of Eagle et al (2002). In keeping with recommended practice for economic evaluations, a condition specific quality of life instrument (the MRF28) and a generic multi-attribute utility instrument (the AQoL) were used to measure health related quality of life. The small scale study that was undertaken by the author to attempt to measure the impact of LTMV on HRQoL in individuals with DMD failed to meet the minimum sample size requirements for the two instruments and consequently it was not able to detect a statistically significant effect.

It should be noted that, given the limited evidence for the effectiveness of the intervention, the inability of the small scale study to recruit and retain a statistically significant sample and the use of a researcher developed model of LTMV service delivery, that the results of the economic evaluation cannot be seen to be anything other than exploratory in nature. In the base-case of the Cost-effectiveness analysis (all costs and outcomes discounted by 3%) Intervention 1 and 2 were dominated by the Current practice
comparator. In the base-case analysis there was a decrease in condition specific quality HRQoL (MRF28) of 28.81% (n=3) following the initiation of ventilation. As the early initiation of LTMV in Intervention 1 and 2 does not provide any additional survival benefit above that gained from the use of the intervention in the Current practice comparator, the event pathways with the shortest duration of ventilation (the Current practice comparator and Intervention 2) with the smallest decrease in mean total condition specific HRQoL were those that were the most cost-effective. In the probabilistic sensitivity analysis, for a maximum acceptable ratio of 100,000 a life-year, the likelihood that the Current practice comparator would be cost-effective was approximately 0.50 to 0.74; for Intervention 2 it was approximately 0.28 to 0.33 and; for Intervention 1 it was 0.04 to 0.18.

In the base-case of the Cost-Utility analysis (all costs and outcomes discounted by 3%) Intervention 1 and 2 produced only small incremental increases in the number of QALYs and there no strategies that were clearly dominated by any other, in either standard or extended dominance. The small incremental gain in the two interventions was due to a 21.89% (n=3) improvement in generic utility weighted HRQoL (AQuoL) after the initiation of LTMV. In the sensitivity and uncertainty analysis the results of the base-case analysis were shown to be just one of a range of possible outcomes, with the Current practice comparator, Intervention 1 and Intervention 2 all becoming dominant for different values of the model parameters. In the probabilistic sensitivity analysis, for a maximum acceptable ratio of $100,000 a QALY, the likelihood that the Current practice comparator would be cost-effective was approximately 0.35 to 0.52; for Intervention 2 it was approximately 0.29 to 0.33 and; for Intervention 1 it was approximately 0.17 to 0.42.

As previously stated, the early initiation of LTMV in Intervention 1 and 2 provides no additional survival and consequently for these interventions to be cost-effective there needs to substantial increase in post-ventilation HRQoL, to offset the costs associated with the increase in the duration of ventilation and the number of ventilated individuals. Although given the severity of the disease, the lack of a cure or any other intervention that significantly prolongs the life-expectancy of individuals with DMD, the standard criteria for cost-effectiveness may create a situation of double jeopardy and the raise the issue of the rule of rescue.

As previously stated, the limited evidence for the effectiveness of the intervention, the inability of the small scale study to recruit and retain a statistically significant sample and the use of a researcher developed model of LTMV service delivery, mean that the results of the economic evaluation cannot be seen to be anything other than exploratory in nature.


# Appendix 1: Respiratory volumes and capacities

<table>
<thead>
<tr>
<th>Measure</th>
<th>Definition</th>
<th>Standard value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Static lung volumes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tidal volume (TV)</td>
<td>Volume of air moved into and out of the lungs with each breath during normal breathing</td>
<td>500 mls</td>
</tr>
<tr>
<td>Inspiratory reserve volume</td>
<td>Volume of air that can be forcefully inhaled after a normal tidal volume breath</td>
<td>3100 mls</td>
</tr>
<tr>
<td>Expiratory reserve volume</td>
<td>Volume of air that can be forcefully exhaled normally after a normal tidal volume exhalation</td>
<td>1200 mls</td>
</tr>
<tr>
<td>Expiratory reserve volume</td>
<td>Volume of air that can be forcefully exhaled normally after a normal tidal volume exhalation</td>
<td>1200 mls</td>
</tr>
<tr>
<td>Residual volume</td>
<td>Volume of air that remains in the lungs after a forced exhalation</td>
<td>1200 mls</td>
</tr>
<tr>
<td><strong>Static lung capacities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total lung capacity (TLC)</td>
<td>Maximum amount of air contained in the lungs after a maximal inspiratory effort</td>
<td>6000 mls</td>
</tr>
<tr>
<td>Vital capacity (VC)</td>
<td>Maximum amount of air that can be expired after a maximum inspiratory effort</td>
<td>4800 mls</td>
</tr>
<tr>
<td>Inspiratory capacity</td>
<td>Maximum amount of air that can be inspired after a normal expiration</td>
<td>3600 mls</td>
</tr>
<tr>
<td>Functional residual capacity</td>
<td>Amount of air in the lungs after a normal tidal volume expiration</td>
<td>2400 mls</td>
</tr>
</tbody>
</table>

Appendix 2: MRF28

INSTRUCTIONS: The following questionnaire asks some questions about your breathing problems, your daily life activities, your attitudes and beliefs. Please try to answer each question and ask if you do not understand anything. Do not think too long about the precise meaning of the question: there are no right or wrong answers.

The following statements are about some daily activities. Answer indicating if these activities usually make you feel breathless. For each item, please tick either true or false as it applies to you.

<table>
<thead>
<tr>
<th></th>
<th>TRUE</th>
<th>FALSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Washing myself (face, neck, etc.)</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>Combing my hair or shaving myself</td>
<td></td>
</tr>
<tr>
<td>1.3</td>
<td>Getting dressed</td>
<td></td>
</tr>
</tbody>
</table>

The following statements are about your daily life and your relations with others (relatives, friends, acquaintances). Indicate if they apply to you because of your breathing problems. For each item, please tick either true or false as it applies to you.

<table>
<thead>
<tr>
<th></th>
<th>TRUE</th>
<th>FALSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>I am unable to shower as I would like to</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>I cannot put on my socks, or shoes as I would like to</td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td>I am not able to cook as I would like to</td>
<td></td>
</tr>
<tr>
<td>2.4</td>
<td>I cannot do things around the house as I would like to</td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>When I need to, I cannot bend over as I would like to</td>
<td></td>
</tr>
<tr>
<td>2.6</td>
<td>I cannot pick up light things as I would like to</td>
<td></td>
</tr>
<tr>
<td>2.7</td>
<td>I cannot play with my friends as I would like to</td>
<td></td>
</tr>
<tr>
<td>2.8</td>
<td>I cannot talk as much as I would like to</td>
<td></td>
</tr>
</tbody>
</table>
Please read the following statements carefully and decide if they give an accurate description of the way you **usually feel these days**. For each item, please tick either true or false as it applies to you.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>I forget names now more than I used to</td>
</tr>
<tr>
<td>3.2</td>
<td>I am very absent minded</td>
</tr>
<tr>
<td>3.3</td>
<td>When I am talking, I often forget what I wanted to say</td>
</tr>
<tr>
<td>3.4</td>
<td>Even when something interests me very much I cannot maintain concentration as long as I would like to</td>
</tr>
</tbody>
</table>

The following statements are about some possible effects caused by breathing problems. Please read them carefully and indicate if they usually apply to you because of your breathing problems. For each item, please tick either true or false as it applies to you.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>I have become an invalid</td>
</tr>
<tr>
<td>4.2</td>
<td>Everything seems too much of an effort</td>
</tr>
<tr>
<td>4.3</td>
<td>I go out to see friends or acquaintances less than usual</td>
</tr>
<tr>
<td>4.4</td>
<td>I spend much more time alone</td>
</tr>
<tr>
<td>4.5</td>
<td>When I am outside I feel I need to have someone with me</td>
</tr>
</tbody>
</table>

The following statements are about your feelings and daily activities. Please indicate if they usually apply to you because of your breathing problems. For each item, please tick either true or false as it applies to you.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>I feel tired in the morning</td>
</tr>
<tr>
<td>5.2</td>
<td>I feel unrefreshed in the morning</td>
</tr>
<tr>
<td>5.3</td>
<td>I feel irritable during the daytime</td>
</tr>
<tr>
<td>5.4</td>
<td>I think my breathing problems are incurable</td>
</tr>
<tr>
<td>5.5</td>
<td>Because of my breathing problems I feel that I am a burden to my family</td>
</tr>
<tr>
<td>5.6</td>
<td>Because of my breathing problems I avoid going shopping</td>
</tr>
<tr>
<td>5.7</td>
<td>Standing up makes me breathless</td>
</tr>
</tbody>
</table>
The following statement is about your **ventilator**. Please tick either true or false as it applies to you. Please tick **NA** only if you do not have a ventilator.

<table>
<thead>
<tr>
<th>TRUE</th>
<th>FALSE</th>
<th>NOT APPLICABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6.1 My ventilator interferes with my life a lot

**HOW IS YOUR HEALTH THESE DAYS?**

TICK ☑ ONE BOX for each question.

7.1 **Overall (general) health:**

| Very good | | |
| Good | | |
| Satisfactory | | |
| Poor | | |
| Very poor | | |

7.2 **Breathing ability:**

| Very good | | |
| Good | | |
| Satisfactory | | |
| Poor | | |
| Very poor | | |

THANK YOU FOR YOUR TIME, YOU HAVE COMPLETED THIS QUESTIONNAIRE
Appendix 3: AQoL1

TICK ☑ ONE BOX for each question to show which statement best describes you during the last week.

1. Concerning my use of prescribed medicines:
   - I do not or rarely use any medicines at all
   - I use one or two medicinal drugs regularly
   - I need to use three or four medicinal drugs regularly
   - I use five or more medicinal drugs regularly.

2. To what extent do I rely on medicines or a medical aid? (NOT glasses or a hearing aid.) (For example: walking frame, wheelchair, prosthesis etc.)
   - I do not use any medicines and/or medical aids
   - I occasionally use medicines and/or medical aids
   - I regularly use medicines and/or medical aids
   - I have to constantly take medicines or use a medical aid.

3. Do I need regular medical treatment from a doctor or other health professional?
   - I do not need regular medical treatment
   - although I have some regular medical treatment, I am not dependent on this
   - I am dependent on having regular medical treatment
   - my life is dependent upon regular medical treatment.

4. Do I need any help looking after myself?
   - I need no help at all
   - occasionally I need some help with personal care tasks
   - I need help with the more difficult personal care tasks
   - I need daily help with most or all personal care tasks.

5. When doing household tasks: (For example: preparing food, gardening, using the video recorder, radio, telephone or washing the car.)
   - I need no help at all
   - occasionally I need some help with household tasks
   - I need help with the more difficult household tasks
   - I need daily help with most or all household tasks.

6. Thinking about how easily I can get around my home and community:
   - I get around my home and community by myself without any difficulty
   - I find it difficult to get around my home and community by myself
   - I cannot get around the community by myself, but I can get around my home with some difficulty
   - I cannot get around either the community or my home by myself.
7. Because of my health, my relationships (for example: with my friends, partner or parents) generally:
   - are very close and warm
   - are sometimes close and warm
   - are seldom close and warm
   - I have no close and warm relationships.

8. Thinking about my relationship with other people:
   - I have plenty of friends, and am never lonely
   - although I have friends, I am occasionally lonely
   - I have some friends, but am often lonely for company
   - I am socially isolated and feel lonely.

9. Thinking about my health and my relationship with my family:
   - my role in the family is unaffected by my health
   - there are some parts of my family role I cannot carry out
   - there are many parts of my family role I cannot carry out
   - I cannot carry out any part of my family role.

10. Thinking about my vision, including when using my glasses or contact lenses if needed:
    - I see normally
    - I have some difficulty focusing on things, or I do not see them sharply. For example: small print, a newspaper or seeing objects in the distance.
    - I have a lot of difficulty seeing things. My vision is blurred. For example: I can see just enough to get by with.
    - I only see general shapes, or am blind. For example: I need a guide to move around.

11. Thinking about my hearing, including using my hearing aid if needed:
    - I hear normally
    - I have some difficulty hearing or I do not hear clearly. For example: I ask people to speak up, or turn up the TV or radio volume.
    - I have difficulty hearing things clearly. For example: Often I do not understand what is said. I usually do not take part in conversations because I cannot hear what is said.
    - I hear very little indeed. For example: I cannot fully understand loud voices speaking directly to me.

12. When I communicate with others: (For example: by talking, listening, writing or signing.)
    - I have no trouble speaking to them or understanding what they are saying
    - I have some difficulty being understood by people who do not know me. I have no trouble understanding what others are saying to me.
    - I am only understood by people who know me well. I have great trouble understanding what others are saying to me.
    - I cannot adequately communicate with others.
13. Thinking about how I sleep:
   - [ ] I am able to sleep without difficulty most of the time
   - [ ] My sleep is interrupted some of the time, but I am usually able to go back to sleep without difficulty
   - [ ] My sleep is interrupted most nights, but I am usually able to go back to sleep without difficulty
   - [ ] I sleep in short bursts only. I am awake most of the night.

14. Thinking about how I generally feel:
   - [ ] I do not feel anxious, worried or depressed
   - [ ] I am slightly anxious, worried or depressed
   - [ ] I feel moderately anxious, worried or depressed
   - [ ] I am extremely anxious, worried or depressed.

15. How much pain or discomfort do I experience:
   - [ ] none at all
   - [ ] I have moderate pain
   - [ ] I suffer from severe pain
   - [ ] I suffer unbearable pain.
Appendix 4: Use of ventilation questionnaire

Use of home ventilation

This form only needs to be completed if non-invasive ventilation is currently being used at home

Today’s Date ………/………/……..

(1) At what age was home non-invasive mask ventilation started? ………… (years)

(2) Why was it started?
(Please tick as many boxes as needed)

<table>
<thead>
<tr>
<th>Problems sleeping</th>
<th>( )</th>
<th>Problems breathing overnight</th>
<th>( )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problems breathing during the day</td>
<td>( )</td>
<td>Problems with chest infections</td>
<td>( )</td>
</tr>
</tbody>
</table>

(3) Over the last month, when has the ventilator been used?
(Please tick as many boxes as needed or provide details)

<table>
<thead>
<tr>
<th>Only when awake</th>
<th>( )</th>
<th>Only asleep</th>
<th>( )</th>
</tr>
</thead>
<tbody>
<tr>
<td>When asleep and at times while awake</td>
<td>( )</td>
<td>Most of the time (up to 16 hours a day)</td>
<td>( )</td>
</tr>
<tr>
<td>All of the time (16 to 24 hours a day)</td>
<td>( )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Other ...........................................................................................................................................
............................................................................................................................................................

(4) Is any form of assisted coughing used at home?
(Please tick as many boxes as needed or provide details)

| No | ( ) | Manually assisted coughing | ( ) |
| Mechanical In-Exsufflation | ( ) | Other | ( ) |

If other, please provide details ........................................................................................................
............................................................................................................................................................
............................................................................................................................................................

Thank you for your time, you have completed this form
Author/s: 
Colgan, Stephen James

Title: 
Economic evaluation of domiciliary long-term non-invasive positive pressure ventilation in Duchenne muscular dystrophy

Date: 
2015

Persistent Link: 
http://hdl.handle.net/11343/58934

File Description: 
Economic evaluation of domiciliary long-term non-invasive positive pressure ventilation in Duchenne muscular dystrophy