Positive Airway Pressure for Obstructive Sleep Apnoea: Systematic Evaluation Versus Clinical and Technological Drift

A thesis presented in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY
in
PUBLIC HEALTH

At Massey University, Wellington Campus,
New Zealand

2005

ABSTRACT

The practice of sleep medicine is expanding and evolving rapidly, often ahead of the evidence base to support clinical practice. Obstructive Sleep Disordered Breathing (SDB) is a condition characterised by repetitive airway collapse causing harmful intermittent blood oxygen desaturations and fragmented sleep. When combined with daytime sleepiness it is known as Obstructive Sleep Apnoea Syndrome (OSAS). Continuous Positive Airway Pressure (CPAP) eliminates SDB by pneumatically splinting open the airway with positive air pressure applied through the nose and/or mouth. CPAP effectively reduces daytime sleepiness in patients with severe OSAS. However, doubt remains as to the effectiveness of CPAP in the majority of patients with mild-moderate OSAS.

The effects of CPAP were compared to a placebo CPAP during a three week crossover Randomised Controlled Trial (RCT) that included 31 mild-moderate OSAS patients. CPAP effectively eliminated SDB (when worn) and moderately improved subjective sleepiness. But, it did not improve objective wakefulness, mood, psychomotor function, or quality of life. Patients who were extremely sleepy at baseline tended to gain the most placebo adjusted benefit from treatment.

A systematic review and meta-analysis aimed to gather and objectively combine all relevant RCT studies to find our whether CPAP reduced sleepiness in patients with mild-moderate OSAS. Seven trials were combined and showed that both subjective sleepiness and objective wakefulness were slightly improved by CPAP therapy.

Objective sleepiness was not improved by CPAP. It is not clear from these two studies that treating mild-moderate OSAS with CPAP is an effective use of resources.

CPAP effectiveness might be limited by sub-optimal compliance. C-Flex aims to improve compliance by modulating pressure during exhalation. C-Flex was compared to CPAP during a pilot RCT that included 19 patients with severe OSAS. C-Flex was associated with a non-significant increase in compliance of 1.7 hours/night compared to CPAP. However, this increase in compliance was not associated with better daytime patient outcomes. Further experiments are proposed as a result of our pilot RCT.

This thesis helps expand evidence-based sleep medicine. Practitioners need to be vigilant, ensuring that treatments are effective in the patients groups in which they are being used (clinical drift), and that new treatments are not adopted without superiority over existing treatments (technological drift).

ACKNOWLEDGMENTS

I have tried to be very clear in the text of this thesis when I'm reporting the products of collaborators' work exactly who they are and what I am grateful to them for providing. But I think it's important to be clear at the start that these studies are the work of a number of people over a number of years and that these contributions require specific acknowledgement.

For the study reported in Chapter 2 I owe a great deal of thanks to my co-authors Deidre Sheppard, Dr Angela Campbell and Dr Alister Neill from WellSleep, Wellington School of Medicine and Health Sciences. Funding for this study was provided by the Health Research Council of New Zealand to Dr Neill. The study conception and design was originally his and has also aided me greatly in being my second supervisor and as a Sleep Physician the prime example of who this information is primarily aimed at. Deidre Sheppard was responsible for the day-to-day running of the study. Gordon Purdie from the Department of Public Health at the Wellington School of Medicine and Health Sciences provided statistical advice regarding mixed models. All polysomnography was collected and scored by Dr Angela Campbell and the technicians at Wellsleep, Karyn O'Keeffe, Margo van den Berg, and Michinobu Imazu. The later stages of this project were also supported by a PhD stipend from the office of the Assistant Vice-Chancellor (Research), Professor Nigel Long. Thirty-one patients gave up their time to help us and were subjected to the PVT and MWT in return. We remain amazed at how much people will put up with to help us. Thank you.

The study reported in Chapter 3 and 4 was also supported by the funding of a PhD stipend from Professor Nigel Long's office and from Professor Philippa Gander

through the Sleep Wake Research Centre. Dr Maree Barnes from the Institute for Breathing and Sleep at the Austin Hospital in Melbourne has been half of the joint effort to produce the systematic review and extract data. Noemie Travier ran the meta-analyses based on the data I had extracted from the manuscripts. This chapter has been submitted to a peer reviewed journal, I am the first author, Dr Barnes the second and Mdme Travier the third. Other authors are Drs Angela Campbell and Alister Neill from the Wellington School of Medicine, Professor Robert Pierce of the University of Melbourne, Professor Doug McEvoy of the Adelaide Institute of Sleep Health and Professor Philippa Gander of the Sleep/Wake Research Centre.

The study reported in Chapter 4 was made possible by the loan of six C-Flex machines from Care Medical, the New Zealand suppliers of Respironics CPAP machines. All polysomnography was collected and scored by Dr Angela Campbell and the technicians at Wellsleep, Karyn O'Keeffe, Deidre Sheppard, Michinobu Imazu, Sue Garret, and Helen Morgan. CPAP compliance was also monitored and collected by WellSleep staff as I was blinded throughout the trial. Twenty-six patients were subjected to the PVT and MWT. Co-authors on the paper under review describing these results are Drs Angela Campbell and Alister Neill.

All of my writings now owe part of their character to my supervisor, Professor Philippa Gander, who has improved significantly all my attempts at communication. Writing about science to her has fundamentally changed the way I think about things and the depth that I think about things. Beneficially this has also made me realize the depths of my own ignorance about almost everything. It's difficult to imagine a better supervisor and I am deeply grateful.

About a year and a half ago at the Sleep/Wake Research centre I got the palpable feeling that we, all, were really getting somewhere. Projects were being finished and new ideas for future projects were being discussed. It has felt good to work with people who know what they don't know (and more importantly know what nobody knows) and are actively working to alleviate that condition. Allison Clark, Riz Firestone, Jesse Gale, Dr Sandy Garden, Dr John Matthewson, Dr Kara Mihaere, Sarah-Jane Paine, Heather Purnell, Dr Leigh Signal, Denise Ratieta, Noemi Travier, and Margo van den Berg: everybody has helped in some way and they have been very good people to be around. Even when they are stressed they seem to remain unflappable. If only all work environments were this good. Dr Kara Mihaere has additionally been extremely helpful in the final stages of preparation in navigating me past the potential formatting minefield of documents this size. Thank you.

To my family: My Mother and Father, Gered and Ray, have always supported my brother and I in whatever we decided to do. As long as we did something and we enjoyed it. They also stopped me from starving during the first year of my PhD before the shift to Massey University (and a stipend!) saved them from having the son that wouldn't leave home. Dr Delwyn Bartlett, my aunt, introduced me to the interesting world of sleep research about 7 years ago, while she was completing her PhD. After that, accountancy didn't seem quite so appealing. Thank you.

My partner, Georgia Foster has caught me on numerous occasions talking about statistics in my sleep. Poor woman. Really don't know how she puts up with me. I'll cook dinner all of next week.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

GLOSSARY & ABBREVIATIONS

AHI	Apnoea/Hypopnoea Index- The most common measure of Obstructive Sleep Apnoea severity. 0-5 /hr Sub Clinical Severity 5-15 /hr Mild 15-30 /hr Moderate 30+ Severe
AutoPAP	Automatically self-titrating continuous positive airway pressure device. CPAP devices that are designed to change pressure in response to the presence or absence of apnoeas, hypopnoeas and/or upper airway resistance. May not be as effective as marketed.
BMI	Body Mass Index- Standard measure of Body Density Weight(kg)/Height^2(m^2). < 20 Underweight 20-25 Normal 25-30 Overweight 30+ Obese
BP	Blood Pressure
COPD	Chronic Obstructive Pulmonary Disease. Also known as chronic obstructive respiratory disease (CORD)
EBM	Evidence based medicine. Application of rationality and scientific methodologies to the provision of healthcare
ECG	Electrocardiogram measurement of electrical output of heart.
EEG	Electroencephalogram measurement of electrical output from brain.
EMG	Electromyogram measurement of electrical output of muscles.
EOG	Electrooculogram: Electrophysiological measurement of eye movements. Standard Measure during polysomnography and during the MWT.
ES	Effect Size. The size of the effect on a given measure divided by the background standard deviation of that measure. Gives the magnitude of the effect which can then be compared between different measures. <0.2 Insignificant 0.2-0.5 Small 0.5-0.8 Moderate >0.8 Large
ESS	Epworth Sleepiness Scale- Most commonly used measure of chronic subjective sleepiness in clinical settings.
CPAP	Continuous Positive Airway Pressure, common treatment for OSAS.

CT	Conservative medical treatment of OSAS.
EDS	Excessive Daytime Sleepiness.
FOSQ	Functional Outcomes Sleep Questionnaire: Sleepiness Related Quality of Life Measure.
GHQ	General Health Questionnaire.
HADS	Hospital Anxiety and Depression Score.
LTSA	Land Transport Safety Authority.
ModMWT	Modified version of the Maintenance of Wakefulness Test. See MWT
mmHG	Millimetres of Mercury. Standard unit measure of blood pressure.
MSLT	Multiple Sleep Latency Test, EEG test of daytime sleepiness. Measures time to fall asleep in a soporific environment when the patient is attempting to fall asleep.
MWT	Maintenance of Wakefulness Test, EEG test of daytime wakefulness. Measures time to fall asleep in a soporific environment when the patient is attempting to stay awake.
MVA	Motor Vehicle Accident
NHP	Nottingham Health Profile.
OA or OD	Oral Appliance or Oral Device. Mandibular advancement splints or tongue stabilisers aimed at increasing airway calibre to treat SDB.
OSAS	Obstructive Sleep Apnoea Syndrome.
OSLER	Oxford Sleep Resistance test. A non EEG alternative to the Maintenance of wakefulness test. Has been shown to be sensitive to treatment for OSAS
PSG	Polysomnography- literally many measurements of sleep. electrophysiological measurements of sleep include EEG, EOG, EMG, ECG. Respiratory channels are also included in the diagnosis and treatment of sleep disordered breathing
RCT	Randomised Controlled Trial.
RDI	Respiratory Disturbance Index: Common alternative but similar measurement to the AHI.
REM	Rapid Eye Movement Sleep also known as Paradoxical Sleep. The EEG is relatively active during REM, the eyes exhibit regular large

	movements but the EMG amplitude is much reduces compared to sleep. Outwardly the person appears very still, but may twitch at the extremities. Respiration and heart rate become more variable than during normal sleep
SaO_2	Percentage of oxygen saturation in blood.
SD	Standard Deviation, statistical measure of spread.
SDB	Obstructive Sleep Disordered Breathing: In this thesis SDB refers to frank apnoeas and hypopnoeas and not to the milder manifestations of obstructive SDB such as upper airway resistance syndrome or snoring.
SE or SEM	Standard Error of the Mean: Statistical measure of precision of an average value
SF-36	Medical Outcome Survey Short Form 36. Common pencil and paper measure of health related quality of life
SHHS	Sleep Heart Health Study: Group of longitudinal cohorts that represent best ongoing investigation into the effects of SDB.
SWS	Slow Wave, or Deep Sleep. Stages 3 and 4 of sleep marked by emergence and then predominance of delta waves in the EEG (0-2 cycles per second). Excludes lighter stages of sleep (stages 1 and 2) and also the qualitatively different REM stage.
UK	United Kingdom.
UPPP	Uvulopalatophyrangoplasty -an operation designed to treat OSAS by ablation of soft tissue in the back of the mouth.

CHAPTER 1
INTRODUCTION: EVIDENCE-BASED SLEEP MEDICINE

1.1 OBSTRUCTIVE SLEEP APNOEA SYNDROME AND SLEEP DISORDERED BREATHING

Introduction

Obstructive Sleep Apnoea Syndrome (OSAS) is a breathing disorder where reversible airway obstruction routinely occurs during sleep. People with this condition often report clinically with loud sonorous snoring and witnessed breathing pauses terminating with gasps, choking sounds and brief arousals from sleep. The immediate cause for seeking medical help has traditionally been daytime sleepiness and/or cognitive difficulties. The clinical presentations of the syndrome have traditionally been described as follows: the patient is male, obese, snores loudly, and is observed to stop breathing during his sleep (usually by bed partner). The patient is also sleepy[1] and may fall asleep in the waiting room where he will begin to snore loudly. However, this characterisation is probably changing, as awareness of the syndrome increases, and bed partners and medical professionals become aware of the potential seriousness of the breathing pauses alone (i.e. without excessive sleepiness). This is largely because epidemiological cross-sectional studies have given a much different description of who has sleep breathing disorders in the community.

It is important to differentiate between Obstructive Sleep Apnoea Syndrome (OSAS) and the larger group of people with Obstructive Sleep Disordered Breathing (hereafter referred to as simply Sleep Disordered Breathing or SDB). OSAS by definition includes SDB plus daytime impairment in the form of symptomatic sleepiness[2]. Population-based studies have identified large percentages of the middle

aged population that have obstructive respiratory disturbance during sleep at pathological levels, and yet seem to be unaffected as far as their daytime sleepiness levels are concerned[3]. Approximately 4% of middle-aged males will have OSAS, but 25% will have SDB[3, 4]. Associations between SDB and hypertension, as well as other cardiovascular diseases, are independent of any daytime sleepiness or other known risk factors[5-10]. Chronic sleepiness can also exist independently of sleep disordered breathing and is found in about 15% of the New Zealand population[11]. These observations have led some commentators to question whether sleepiness in many OSAS patients is really caused by SDB, by chance association, or by a third factor that might cause them both[12, 13]. Cross-sectional and longitudinal cohort data points towards the SDB being the driving force behind the exacerbated risks of hypertension and cardiovascular disease, and not the additional sleepiness that qualifies for OSAS. Large numbers of people in the community have SDB and are unlikely to ever be treated, given the current practice of treating the hyper-somnolent, who self present. Partially, this practice is due to there being no demonstrably effective treatment for this non-sleepy group.

Clinical Definition

Obstructive sleep apnea syndrome *is characterised by repetitive episodes of upper airway obstruction that occur during sleep, usually associated with a reduction in blood oxygen saturation.[2].*

Diagnostic Criteria: Obstructive Sleep Apnea Syndrome (780.53-0 ICD-9-CM)[2]

- ***A.*** *The patient has a complaint of excessive daytime sleepiness or insomnia. Occasionally the patient may be unaware of clinical features observed by others.*
- ***B.*** *Frequent episodes of obstructed breathing occur during sleep*

C. *Associated features include:*

a. *Loud Snoring*

b. *Morning headaches*

c. *A dry mouth upon awakening*

d. *Chest retraction during sleep in young children*

D. *Polysomnographic monitoring demonstrates :*

a. *More than five obstructive apneas greater than 10 seconds in duration, per hour of sleep and one or more of the following*

i. *Frequent arousals from sleep associated with the apneas*

ii. *Bradytachycardia*

iii. *Arterial oxygen desaturation in association with the apneic episodes*

b. *Multiple Sleep Latency Test may or may not demonstrate a mean sleep latency of less than 10 minutes.*

E. *The symptoms can be associated with other medical disorders (e.g. tonsillar enlargement).*

F. *Other sleep disorders can be present (e.g. periodic limb movement disorder or narcolepsy).*

Minimal Criteria: *A plus B plus C.*

In adults, an apnoea is a cessation of breathing for ten seconds or more, while an hypopnoea is a substantial reduction in breathing (50-90% reduction) for 10 seconds or more. Apnoeas and Hypopnoeas are of three different types:

- Inhalation attempts that are hindered by the collapsed airway are termed obstructive events (i.e. abdominal breathing effort but without filling of the lungs evinced by little thoracic expansion).
- When inhalation is not attempted, these events (which may or may not include an occluded airway) are termed Central Apnoeas (i.e. lack of both thoracic expansion and abdominal effort).
- The third type of apnoea/hypopnoea is a combination of effort and no effort with insufficient thoracic extension and is called a Mixed Apnoea.

In population based estimates, by far the most common type observed is the obstructive apnoea or hypopnoea[3]. If the events are mostly central in nature, then the diagnosis should be Central Sleep Apnoea. This thesis focuses on the treatment of the obstructive type of apnoea/hypopnoea.

The numbers of apnoeas and hypopnoeas are usually summed and divided by the number of hours of sleep to give the main index of disease severity: the apnoea hypopnoea index (AHI). AHI scores below 5 per hour are usually regarded as of subclinical severity, 5-15/hr as mild SDB, 15-30 moderate, and above 30 as severe[14].

Causes of Airway Collapse

Humans are almost uniquely affected by obstructive SDB. The only other reported animal to have an SDB-like syndrome is the English Bulldog, but this phenomenon is mostly restricted to the animal's REM sleep. The English Bulldog has therefore been an animal model for human SDB[15, 16]. One proposed reason that SDB is unique to humans relates to our airway, which is uniquely adapted. Recent expert reviews of airway obstruction during sleep present an uncertain, but almost certainly multifactorial causal sequence[17, 18].

Humans have a free-floating hyoid bone that, unlike other mammals, is not directly attached to any other bone or cartilage. Muscular contraction or relaxation whilst supine (and the weight of any fat and/or muscle overlying this area) is not held in check by a structural anchor point, and can therefore increase outside pressure on the upper airway. Sforza and colleagues[19] investigated the pharyngeal collapse in severe SDB sufferers and found that the position of the hyoid bone, and abnormally large amounts and specific sites of soft tissue in the pharynx, contributed to airway collapse.

The second and third reasons are closely related to the effects of air movement in the human pharynx, which has multiple functions including respiration, speech and swallowing[17]. It is long, compared to that of other mammals, and has an almost ninety degree turn at the level of the nasopharynx. These two factors affect the airflow through the pharynx and are hypothesised to increase the potential for airway collapse.

While a person with SDB is awake, the patency of the airway is maintained by the pharyngeal dilator muscles. If these fail during sleep (with or without a structural abnormality) the airway may become totally (apnoea) or partially (hypopnoea) occluded. Occlusion usually happens at the end of exhalation, just as inhalation begins and where the lungs are relatively empty.

During sleep in healthy people, it is possible to mimic moderate SDB by using 10 mmH_20 of negative pressure, which led King and colleagues[20] to conclude that decreased pharyngeal pressure alone can be responsible for airway collapse in healthy subjects. Interestingly, upper airway muscular tone in patients with OSAS is about twice as high during wakefulness than in people without SDB. This is possibly a reaction to increased resistance in the airway through narrowing or obstruction that is being compensated for during wake, but not sleep[21].

It has been suggested that the tendency for OSAS to run in families may be due to poor hypoxic ventilation response[22, 23]. This pattern within families may also suggest some sort of inherited susceptibility toward airway narrowing or collapse[23].

Many people with SDB have structural abnormalities in the upper airway. In the case of sudden onset of SDB, the development of an airway abnormality may be a direct causative factor. For instance Walshe and colleagues[24] have presented a rare case of SDB in a 72 year old man being caused by a superior laryngeal schwannoma. The schwannoma is this case was causing a ball valve like obstruction of the airway during inspiration. Removal of the schwannoma resolved the patient's sleep disordered breathing. Structural abnormalities are, however, not necessarily always present in those with SDB[25].

Impaired afferent sensation in the upper airway may also play a role in obstruction. Upper airway vibration may cause impaired sensation in people with SDB and/or snoring that might leading to apnoeas of longer duration[26]. Gora and colleagues[27] compared people with mild OSAS to age and BMI matched controls and found that the OSAS patients had a blunted cortical response to respiratory occlusion, but this response difference was not observed when the subjects were awake. Upper airway vibration may also cause upper airway oedema, which further exacerbates the collapse of a relaxed upper airway. Alternatively the observed level of inflammation in OSAS, patients may have been caused by systemic inflammation associated with OSAS rather than direct vibration of the upper airway through snoring[28].

Thus, multiple factors can result in a pressure imbalance (largely structural causes) which is not counteracted by the airway dilators (largely neurological causes) as it is during wake[29-31]. SDB in an individual usually has multiple causes, but can sometimes be caused by a single pathology, if that pathology is severe enough.

1.2 THE EPIDEMIOLOGY OF SLEEP DISORDERED BREATHING

Population Prevalence

Cross sectional findings from both Busselton in rural Australia and the Wisconsin cohort, indicate that about 25% of middle aged men have an Apnoea Hypopnoea Index>5/hr. Among middle aged women in the Wisconsin cohort, about 9% had an AHI of at least 5/hr. The additional diagnosis of OSAS, which requires SDB and significant daytime sleepiness, was met by 4% of men and 2% of middle aged women. [3, 4, 7]. Major risk factors for sleep disordered breathing in the Sleep Heart Health Study include male gender, increasing age up to about 65 years, obesity, increasing neck size, and loud habitual snoring[32]. In the Bussleton study, MESAM IV* identified that 26% of men had a respiratory disturbance index of greater than 5/hr and 3.4% had an RDI greater than 20/hr. Only obesity (Body Mass Index†) predicted sleep disordered breathing in these 294 men. In relatively affluent Wellington, New Zealand, the picture is similar, with a MESAM IV study identifying 22% of community dwelling Māori men, 11.4% of non-Māori men, 6.3% of Māori women, and 3% of non-Māori women having an RDI≥ 5/hr[33]. New Zealand and Australia also have substantial and growing populations of people of Asian and Polynesian descent for whom prevalence is unknown and risk factors for OSAS might differ significantly. Polynesian peoples, in particular seem likely to have a very high prevalence of SDB[34].

* A four channel overnight monitoring of heart-rate, body position, peripheral blood oxygen saturation via pulse oximetry, and snoring volume

† Body Mass Index is defined as the weight of a person in kilograms divided by the square of their height in metres. In most studies and by general consensus obesity is where BMI has exceeded 30 kg/m^2.

Risk Factors for Sleep Disordered Breathing

Gender

Males are twice as likely as females to have SDB or OSAS in middle age[3, 35, 36]. The disparity between genders is not as marked as was once believed- the preponderance of males over females in some American clinical populations was around 8:1[37]. By comparison recent, New Zealand based RCTs contain around 4 to 8 times as many men as women [38-40].

Clinical reporting bias likely explains much of the difference between the clinical populations and the prevalence estimates from the population studies. Basing referrals on BMI and male gender would miss much of the population affected by OSAS or SDB[32]. Clinical bias means that women might be suffering the downstream effects of SDB disproportionately more than men.

There are a number of plausible mechanisms to explain why men are predisposed to SDB[41]. Magnetic Resonance Imaging (MRI) scans of 10 matched male and female pairs have shown that males have a larger pharangeal soft tissue mass than females[42]. This may predispose men to airway collapse during sleep due to additional weight on the airway. Some commentators have postulated that some of the gender difference may be due to the mechanical properties of the longer and larger male airway, coupled with a larger soft palate[43].

Age

In the Sleep Heart Health Study (SHHS), the proportion of people with SDB (AHI>5) increases steadily from about age 35 (1 in 20), until about age 65, where it

seems to stabilise at about 1 in 5, or less. Older age groups (75 or older) are somewhat less stable in the prevalence estimates available, but this is due to the decreasing sample size in the advanced years[32]. Other studies also indicate that the prevalence of SDB increases through middle age and up to around the age of 65 years when the prevalence stabilises[32, 36, 44]. This might be an artefact of lack of data from these older age groups. It is also consistent with a survival effect, where a significant proportion of people with SDB might have already succumbed to fatal cardiovascular disease (CVD). This observation has recently been given some support from the SHHS where SDB was only a risk factor for various definitions of hypertension in those under the age of 60 years[45].

Ethnicity

Studies of community levels of SDB have often been done in people predominantly of European descent, for example, the Wisconsin[3] and Busselton studies[4], leaving little ability to reliably calculate prevalence or risk factors for OSAS in minority populations. However the Sleep Heart Health Study also contains some cohorts with significant numbers of Native Americans and African Americans. Young has found that African Americans in this study had a 20% prevalence of AHI≥ 15 and that American Indians had a 23% prevalence, compared to the 17% observed in the White population. Once body habitus had been controlled for, neither African Americans or American Indians were more at risk of sleep disordered breathing[32]. In contrast Ancoli Israel *et al.* have found that, even after controlling for BMI and other confounders, African Americans were at 2.5 times the risk of AHI≥ 30 (measured with in-home polysomnography) than Whites in a community dwelling sample of people over 65 years.

New Zealand based studies were needed to estimate prevalence of OSAS in Māori. Māori were oversampled to enable modelling of the risk factors for OSAS/SDB to the same level of precision as in the majority population of European descent (Pakeha). Consequently, the three population based studies in New Zealand (two postal, one using MESAM IV) have had roughly half Māori and half Non-Māori participants[11, 33, 46, 47], even though only 15% of the population identifies themselves as Māori[48]. Māori have approximately twice the prevalence of sleep disordered breathing as Non-Māori in New Zealand[33]. However, this difference disappears after controlling for higher levels of obesity and larger neck circumferences in Māori[33].

Asian populations, particularly ethnic Chinese, are also growing sectors of New Zealand and Australia societies. Cranial morphology in these groups might indicate a different set of risk factors for SDB. Hong Kong based studies have found that men (n=153), aged 30-60, have a 25% prevalence of AHI≥ 15 and a 4% prevalence of OSAS (AHI≥ 5 and excessive daytime sleepiness)[49]. In women (n=106) the pattern was almost identical to western estimates, with 2% of Hong Kong women with OSAS but 10% with at least moderate SDB (AHI≥ 15)[50]. Whilst BMI and other measurements of body habitus were risk factors for SDB in Hong Kong workers, these predictors were not as strong as in White Americans[49]. The overall prevalence of obesity in this Asian population is also lower than for American samples.

Kim *et al.* have investigated the prevalence of OSAS in 5020 Korean men and women aged 40-69 years, using full overnight polysomnography in a subset of 457[51]. The prevalence of SDB (AHI>5), extrapolated back to the full sample, was 27.1% in men and 16.8% in women. OSAS was found in 4.5% of men and 3.2% of women.

Logistic regression analysis found that SDB was associated with sex, body mass, and hypertension. The peak in prevalence seemed to be in the 50's, with 25.2% of women and 33.7% of men having an AHI≥ 5/hr. The prevalence was similar in the 60-69 year old age range, but substantially lower in the 40-49 years decade, at 24.2% in Men and 8.2% in Women.

Body Mass and Obesity

Excess weight is a commonly identified risk factor for SDB. Multiple measures of obesity have been found to predict SDB including neck size, central obesity and general obesity[7].

Longitudinal changes in AHI were associated with changes in weight over 4 years, in 690 men and women from the Wisconsin cohort[52]. A one percent increase in (or decrease) in body weight was associated with a 3% increase (or decrease) in the AHI, and this association fitted well with other studies of weight loss and AHI reduction[7]. Weight gain is almost certainly a causal factor in the development of SDB and thus is a key modifiable risk factor.

In New Zealand the picture is similar, with incremental increases in BMI or neck circumference both found to increase the risk of having sleep disordered breathing[33].

Incidence and Progression

The Wisconsin cohort currently provides the best information on incidence and progression of SDB. However, it has the limitations that the study population is middle class and predominantly mono-cultural. At baseline, 644 participants had sub-clinical or no SDB (AHI<15/hr). At the 4 year follow-up, 39 of these had measured AHI≥15,

giving an incidence of around 6% or 1.5% per annum[52]. The incidence and remission of SDB in this group seems to closely follow weight loss or gain (see above)[52]. The odds of having moderate-severe OSAS (AHI≥15) were 6 times higher after a 10% weight gain[52]. Eight year followup data is also available in a smaller number of participants whose everage AHI has increased from 2.5 to 5.1/hr[7]. Very few people with significant SDB at baseline have improved to where they have very low AHI (AHI≤1) at follow-up. Disentangling the contributions of age and weight gain to SDB remains a difficultly in interpreting these studies[7].

1.3 HEALTH-RELATED CONSEQUENCES OF SLEEP DISORDERED BREATHING AND OBSTRUCTIVE SLEEP APNOEA

Sleepiness

A clear, but weak, dose response relationship exists between measured levels of SDB (the AHI) and excessive daytime sleepiness. Many people with excessive daytime sleepiness do not have SDB and many of those with significant levels of SDB do not experience excessive daytime sleepiness[3].

In the Sleep Heart Health Study, people with an AHI >30 had a mean Epworth Sleepiness Score (ESS)[‡] of 9.3, whilst those with an AHI<5 had a mean Epworth of 7.2. Twenty one percent of people with AHI<5 had an Epworth above 10 (excessively sleepy), whilst 35% of those with AHI>30 had ESS>10[53]. In the Wisconsin cohort, excessive daytime sleepiness was defined as experiencing the combination of these three factors, at least twice a week: waking unrefreshed no matter how long they had slept, plus excessive daytime sleepiness, plus daytime sleepiness that interfered with daily living[3]. Among non-snorers, 10% of women and 3% of men were excessively sleepy. By comparison, among people with AHI>5, 23% of women and 16% of men were excessively sleepy.

In the Cardiovascular Health Study (n=4578, aged ≥ 65), excessive daytime sleepiness measured by the Epworth had the following independent risk factors; non-white race, depression, loud snoring, awakening with dyspnoea or snorting (suggestive of SDB and/or COPD), frequent nocturnal awakenings, medications for congestive

[‡] The Epworth Sleepiness Scale is the most common measure of trait sleepiness in sleep medicine. It is scored out of 24 and scores above 10 are regarded as evidence of excessive daytime sleepiness. The measure is discussed at length in Appendix 1

heart failure, non-use of sleeping pills, sedentary lifestyle, and limitations of daytime activities[54]. Although sleep was not measured objectively, these findings indicate that some of the risk factors for SDB are associated with daytime sleepiness.

In New Zealand, 15% of 30-60 year olds had an Epworth score above 10 [11]. In logistic regression models, with a sample of roughly half Māori and half non-Māori participants, independent predictors of sleepiness were; Māori ethnicity, male gender, being older, having observed apnoeas, being from a poorer neighbourhood, self-reported sleeping outside 6.5 to 8 hours per night, never or rarely getting enough sleep, never or rarely wakening refreshed, snoring, having a large neck, and being a non-drinker. Many of these factors, such as large neck size, male gender, Māori ethnicity, being older, nightly snoring, and having observed apnoeas, are also recognised risk factors for sleep disordered breathing[33].

Hypertension

Possible Causal Mechanisms and Animal Studies

The prevalence of hypertension (defined as 160/95 mmHg) in New Zealand, for those over 45, has been reported as 29% for men and 24% for women[55]. Maori, Pacific Island and Asian populations have slightly increased average blood pressure (BP) compared to European New Zealanders[56]. Body mass index explains about half of the variation between smaller ethnic groups and European New Zealanders in this study, but not in those of Asian descent[56]. Hypertension is well recognised as a modifiable risk factor for the later development of cardiovascular disease and premature death[57] (www.moh.govt.nz). There is increasing evidence that many patients presenting with hypertension also have sleep disordered breathing, and that alleviation

of SDB via CPAP may reduce the risk of cardiovascular disease attributable to hypertension in these patients[58, 59].

Breath holding itself (apnoea) results in an acute rise in blood pressure and heart rate[60, 61]. The forced inspiration against an obstructed airway (Mueller manoeuvre) also acutely increases BP and heart rate[62]. Hypoxemia also leads to significant acute increases in blood pressure[63, 64]. These three phenomena are all present during SDB, resulting in abnormal variability in blood pressure and heart rate during sleep. The resulting sympathetic activation is considered to be a cause of the observed chronic increases in BP.

Earlier studies of the association in humans between SDB and hypertension have been criticised[12, 65, 66]. However, compelling multiple lines of recent evidence give strong support to the claim that SDB causes hypertension[6, 59].

Sleep Disordered Breathing experimentally induced in dogs over a 1-3 month period is associated with chronic increases in both night and daytime blood pressure, by up to 15mm Hg. These elevations disappeared within one month after reversal of SDB[67]. Elevations in BP are known to be caused by arousals from sleep (probably via sympathetic activation[68]), but aurally induced arousals from sleep were shown not to induce chronic rises in BP in these dogs[69].

Fletcher and colleagues[70] have demonstrated that intermittent hypoxia can raise daytime BP by about 12mm Hg in rats. This chronic change is not due to sleep arousals[71], which are known to produce acute rises in BP. Interestingly the development of daytime hypertension is not seen in all genetic strains, leaving open the

intriguing possibility that SDB may induce hypertension only those humans who are genetically susceptible. Increased susceptibility to hypertension induced by SDB might explain why some of the clinical populations studied have shown little increased risk of hypertension with increasing AHI. A survival effect might result in patients in clinical studies being those relatively resistant to the effects of high AHI on BP. This would be a problem where there is a large lag between the development of SDB and the subsequent diagnosis of OSAS.

It thus seems that the effects of hypoxia, hypercapnia, increased pleural pressure, and the increased activation of the sympathetic nervous system, but not the sleep fragmentation associated with SDB, are the likely causes of hypertension in this condition[6, 72].

In addition to the effects of SDB on daytime hypertension, it is notable that nocturnal apnoea and associated respiratory effort and arousals are closely associated with acute spikes in blood pressure during sleep. As the average human sleep time is 7-8 hours/day, these sleep-related events may account for $1/3^{rd}$ of the blood pressure related mortality observed in these patients, even assuming zero effect on blood pressure during wakefulness.

Cross Sectional Studies

Bixler and colleagues[73] have reported that sleep disordered breathing and snoring, together or separately, are independent risk factors for hypertension (n=1741, aged 20-100). There was a linear dose-response relationship between SDB and the risk of hypertension, but interestingly as age increased the strength of association declined. Indeed, in the older age groups, the relationship was inverted. The strongest relationship

between SDB and hypertension was observed in young males with lower BMI and severe SDB.

In the Wisconsin cohort (n=936, aged 30-60 years), a linear increase in BP has been observed with increasing AHI, which was independent of other factors[74]. At the average BMI (30) for this cohort, it was observed that the group with the highest AHI (>15) had 3.6 mmHg higher systolic and 1.8 mmHg higher diastolic pressures than did the group with AHI=0[75]. Participants with AHI≥ 30 were 3.07 times more likely to have hypertension or be on hypertensive medication than the reference group of AHI=0. For a given AHI, as BMI increased, the effect on blood pressure decreased. Furthermore younger, leaner people seem to be at higher additional risk of developing hypertension if they have SDB, once other factors are controlled for. The relationship between SDB and hypertension is not strong, but it is statistically significant[75]. These findings are also based on cross sectional, and not longitudinal data, which would offer stronger evidence for a causal connection[74].

Likewise, cross-sectional findings from the larger Sleep Heart Health Study (n=6132), found a linear relationship between hypertension (140/90 mm Hg or more) and increasing severity of sleep disordered breathing[76]. People with the most severe sleep disordered breathing (AHI ≥30) had an independent odds ratio for hypertension of 1.37 (95% CI 1.03, 1.83), compared to the lowest severity of SDB (AHI<1.5) after controlling for the usual risk factors.

A clinical cross-sectional study[5] has demonstrated an independent linear increase in the risk of hypertension as severity of OSAS increased, independent of other known risk factors (n=2677). An increase of 10 apnoeic events per hour was associated

with an increased risk of about 11% of having hypertension. Compared to the standard severity classifications[14], this would mean that the odds ratio for the clinical severity classifications (AHI=0-4.9, 5-14.9, 15-30, and >30) would be roughly as follows, 1.00 (reference group- less than 5 events), 1.11 (15 events), 1.22 (25 events), 1.33 (35 events). These findings are thus highly comparable to Young *et al.*[74] from the Wisconsin Cohort (ages 30-60 years), where the group who had AHI>30 had an adjusted odds ratio of 1.37 (CI 1.03-1.83) of having hypertension (vs. reference group with AHI<1.5). Lavie's model also calculates that a person with an AHI of 60 will have a mean arterial pressure about 5.5mmHg higher than if they did not have SDB. The same group has also shown that OSAS patients with hypertension have higher mortality than people with OSAS alone[77].

Longitudinal Studies

The Wisconsin Sleep Cohort is now beginning to produce longitudinal comparisons[10]. Sleep disordered breathing is a significant independent and prospective risk factor for developing hypertension (140/90 mmHg or hypertensive medication) within 4 years, after adjusting for baseline hypertension, age, sex, BMI, alcohol and cigarette consumption, neck and waist measurements (see Table 1.1). The investigators also found that, when they regarded AHI as a linear variable, there was no level below which a single unit of AHI did not increase the odds of the later development of hypertension.

Table 1.1: Prospective Relationship Between Sleep Disordered Breathing and Later Development of Hypertension

Baseline AHI	0 (n=187)	0.1-4.9 (n=507)	5.0-14.9 (n=132)	≥ 15.0 (n=67)
Mean Baseline BP	120/79	124/82	130/84	135/88
Mean BP 4 years	118/75	123/79	131/82	129/81
% Hypertensives at baseline (≥140/90 or medication)	18	24	45	60
% Hypertensives at follow-up (≥140/90 or medication)	17	28	48	60
OR for Hypertension	1.0	1.42	2.03	2.89
95% CI for Odds Ratio	REF	1.13-1.78	1.29- 3.17	1.46- 5.64

From Peppard PE, *et al.* Prospective study of the association between sleep-disordered breathing and hypertension. NEJM 2000; 342 (19): 1378-1384.

Treatment Studies

Randomised controlled treatment trials of CPAP are reviewed later[78-83], but one uniquely elegant non-randomised study deserves mention because it shows that the reduction of SDB, and not some other action of CPAP, is responsible for meaningful reductions in blood pressure[58]. The investigators recruited patients from a hypertension clinic and then screened participants for SDB. Those with SDB and hypertension (SDB+HT, n=14) were manually titrated to a CPAP pressure that abolished their SDB. The group with hypertension but no SDB (HT, n=10) were fitted with CPAP machines that were set to deliver 5cmH_2O. Three weeks of CPAP therapy was sufficient to significantly reduce mean systolic (-7.8 mmHg) and diastolic BP (-5.3 mmHg) in the SDB+HT group but not in the HT alone group(0.3 and -0.7 mmHg respectively, p for differences both<0.05). Combined with the animal studies and growing human epidemiological literature, this study lends further strength to a conclusion that hypertension is caused by SDB.

Issues and Implications

Epidemiological evidence is mounting that there is a moderate but significant independent relationship between hypertension and SDB[7, 59]. The main source of these high quality data is from the Wisconsin cohort and the larger Sleep Heart Health Study (which also contains the Wisconsin cohort). Comparisons between these studies are somewhat complicated by the differing methodologies used. These relate particularly to the differing reference groups used for calculating relative risk, and the differing methods for calculating hypopnoeas[84]. Studies such as those by Peppard *et al.*[10] or Young *et al.*[74] use a reference group with a negligible AHI (AHI=0 or AHI<1.5), whereas studies by other groups have used reference groups with low levels of sleep disordered breathing (often AHI<5). This latter approach may have the effect of underestimating the strength of the relationship between hypertension and SDB. The current system for classifying AHI<5 as of subclinical importance[2] is arbitrary, and is being challenged by the findings from the cohort studies. AHI might eventually be seen in a similar light to blood pressure, where any reduction in levels is desirable as it results in a decreased risk of cardiovascular disease.

OSAS is a risk factor for pharmacologically resistant hypertension[85]. In 183 OSAS patients (AHI>10/hr), those patients reporting hypertensive medication use, but having uncontrolled hypertension, had significantly higher AHI (44±29 vs. 33±25 events/hr, $p<0.0005$), after adjustment for age, gender, and BMI. This implies that SDB-induced hypertension is less likely to be amenable to standard therapy, and different treatment may need to be used in patients with both severe OSAS and hypertension.

The true risks may also be underestimated as an unfortunate side-effect of the multiple regression techniques used[6, 72]. Many of the predictors of SDB, such as BMI, are also significant predictors of hypertension. Controlling for these closely related variables may underestimate the true risk, if these factors are true casual factors related through SDB. This is especially true where a pernicious bi-directional relationship is thought to exist between SDB and other recognised CVD risk factors[13]. SDB might cause obesity, as well as be partially dependent upon it, through the action of insulin resistance[86-88]. Both recurrent hypoxia and sleep disruption/restriction, are thought to lead to insulin resistance. When seen in this light, the chronic sleep restriction of whole societies, coupled with a noted rise in obesity, the metabolic syndrome, and SDB, represents a large public health issue[86].

If SDB becomes more widely recognised as a cause of cardiovascular disease, it might be seen as a treatable risk factor. Hypertension, in the presence of SDB, might be remediated with OSAS treatments, rather than standard antihypertensive medications[6, 72]. This rationale has been tested in RCTs which have shown significant improvements in some measures of BP with CPAP in people with OSAS[78-80, 82]. Mandibular advancement splints have also been shown to lower some blood pressure parameters, by treating underlying SDB[89, 90]. Treating hypertension in the larger group of people with SDB (and no daytime sleepiness) has not yet been tested in an RCT, except by Barbe *et al.* who found no effect[83]. CPAP seems unlikely to be a long term treatment option for people with asymptomatic SDB[83].

Heart Failure

Cross-sectional analyses of the Sleep Heart Health Study have found that participants with an AHI>11/hr had 2.2 times the risk of heart failure compared to those with AHI>1.4/hr, and that a significant linear trend existed between AHI and the risk of heart failure[9]. In a parsimonious model, which did not control for hypertension-related factors and smoking, those with AHI >11 had a greater risk of heart failure than those with than AHI<1.4 (OR 2.38 95% CI 1.22, 4.62) (see below: Cardiovascular disease in general).

Peker and colleagues[91] report a 7 year follow-up of 60 men with OSAS and 122 men without OSAS/SDB. OSAS status was established at baseline with a mixture of oximetry, chest wall movements and nasal air flow. Those with more than 30 desaturations (4%) in blood oxygen per hour (with confirmation from other measures) were deemed to have OSAS. All participants (with or without OSAS) were free of hypertension, pulmonary disease, cardiovascular disease, type II diabetes, alcohol dependency, or psychiatric disorders at baseline. OSAS at baseline was associated with an odds ratio of 4.9 (95% CI 1.8-13.6) for having CVD 7 years later, after controlling for age, OSAS treatment, body mass index, baseline systolic and diastolic blood pressure, and smoking habits. However, in participants with OSAS, effective treatment was found to be strongly protective against CVD (OR 0.1 95% CI 0.0-0.7), after controlling for age and systolic blood pressure at baseline. In the whole cohort, the strongest predictor of CVD was increasing age (OR 23.4, 95% CI 2.7-197.5).

Stroke

Possible Causal Mechanisms

OSAS is associated with an increased thickness of the common carotid artery walls, a known risk factor for stroke[92]. Twenty-three pairs of people with severe OSAS were matched to people without SDB, and then scanned with high resolution ultrasonography. These indicated that the OSAS patients had significantly thicker arterial walls (1.43±0.34mm for OSAS and 0.976±0.17mm for controls). However, as this is a case control study, arterial thickening of the carotids may be due to some other factor/s. Nevertheless, the study does offer some strength to the argument that SDB might cause stroke, perhaps through the considerable chronic vibration of the adjacent carotid arteries via the action of snoring[93].

Cardiac arrhythmias are also a leading risk factor for stroke. A recent review reports that atrial fibrillation (AF) in particular seems common in people with SDB/OSAS[94], and CPAP has been reported to alleviate AF in patients with OSAS in a non-randomised study[95]. However, the current evidence is of preliminary status and is not yet compelling. Such an association could be explored in the Sleep Heart Health or Wisconsin studies, but the investigators have not yet published the relevant data.

Patients who have had a stroke have been shown to be more likely to have SDB than matched controls[96]. However, this is not necessarily strong evidence that SDB might predispose people to cerebrovascular incidents. Indeed, it is likely that the stroke itself may have caused or exacerbated the observed SDB, or possibly a third factor might cause both SDB and stroke. Longitudinal data in patients with verified SDB who

then develop stroke, would offer better evidence of causation. Some consensus exits that the SDB observed in these types of studies predates the stroke[72], but this in itself is weak evidence of causation. Animal models may offer some evidence that SDB causes stroke. Longitudinal cohort data should be available to address this question within the next 5 years.

Cross Sectional Studies

The Sleep Heart Health Study[9] showed a non-significant linear trend (p=0.06) of increasing risk of stroke as AHI increased. The highest quartile in terms of AHI (>11) had a relative odds ratio of 1.55 of reported stroke compared to the reference quartile (AHI<1.4), however this relationship was not significant (95% CI 0.96, 2.50). Nevertheless, in a model that removed some of the variables that are highly co-correlated with both stroke and SDB, the risk of stroke increased linearly with AHI (p=0.03). In that model the odds ratio of stroke in the highest quartile was 1.58[9].

Cardiovascular Disease in General

Sleep Heart Health Study cross-sectional data also yield information about general cardiovascular diseases (CVD) [9]. SDB was split into rough quartiles of severity in these analyses (Quartile I (reference), AHI=0-1.3, Quartile II, AHI=1.4-4.4, Quartile III, AHI=4.5-11.0 and Quartile IV, AHI>11). Cardiovascular disease was defined as having been told by a doctor that one had angina, heart attack, stroke, or heart failure AND/OR having undergone coronary bypass surgery or angioplasty. After controlling for age, sex, race, BMI, smoking status, diabetes, self-reported hypertension, use of anti-hypertensive medication, systolic BP, total cholesterol, and high density lipoprotein cholesterol, only quartile IV (i.e. AHI greater than 11/hr) showed significant

increased risk of CVD (OR=I.30 95% CI= 1.01-1.67). However, the full model exhibited a significant upward linear trend across quartiles (p=0.01). These authors also employed a 'parsimonious model' which was calculated via a forward selection procedure that excluded 5 variables on the basis that inclusion of these variables may over-control for the effects of SDB (smoking status, BMI, self-reported hypertension, use of hypertensive medication, and systolic BP). This model showed a significant upward linear trend across quartiles (p=0.0003). Participants in quartile III (RO 1.28 95% Cl 1.02-1.61) and quartile IV (RO 1.42 95% C1 1.13-1.78) were at significantly higher risk of CVD, when compared to those in quartile 1[9].

Newman and colleagues[97] report an association between SDB and known CVD risk factors from the Sleep Heart Health Study. In men, age, smoking status, waist to hip measurement, hypertensive status, and low density lipoprotein cholesterol (in aged <65 years only) all increased linearly with AHI. In woman, risk factors additionally included: high density lipoprotein cholesterol (inversely related), and African American ethnicity, but did not include waist-hip measurement, hypertensive status and overall cholesterol. Many of these risk factors for CVD are often found clustered with each other and with SDB[98]. Obesity, for instance, is a worrisome confounder in these relationships, as it might be both a cause and consequence of SDB. Controlling for obesity in these models may underestimate the true strength of the relationship. Failing to control for it almost certainly overestimates the risk.

General Note

The Sleep Heart Health data and the Wisconsin Cohort give interesting information about possible linkages between SDB and hypertension, and any

subsequent link to CVD. Firstly it seems almost certain that SDB is independently associated with hypertension, but that this effect is of moderate size. This conclusion is based on multiple lines of evidence, including epidemiological studies, human and animal physiological studies, and reduced BP indices in some human treatment trials. Hypertension caused by SDB should, in turn, eventually lead to increased risk of CVD and mortality in these cohorts, as they age. However, the link between SDB and heart failure has been shown to hold even when controlling for such factors as hypertension and BMI. This leads to the conclusion that hypertension is not the only agent mediating the relationship between SDB and heart failure. Some other effect, or uncontrolled correlate, of SDB must be at work. Mortality data have been reported at some conferences, but have not yet appeared in the peer-reviewed literature. The general pattern seems to suggest that as SDB severity increases, so too does the risk of mortality, and that the effects of SDB might be particularly marked in women, who are currently under-treated[99].

Two very recently published longitudinal cohorts of patients referred to sleep clinics are instructive[100, 101]. Marin *et al.* performed diagnostic polysomnography on 1651 men and recorded a mixture of patients with no discernable sleep disorder (n=264), simple snoring (n=377), untreated mild-moderate OSAS (n=403), untreated severe OSAS (n=235, refused or were intolerant of CPAP treatment), and CPAP treated severe OSAS (n=372). Untreated severe OSAS was independently associated with higher risk of both fatal and non-fatal cardiovascular incidents when compared to the healthy men (odds ratios 2.87 and 3.17 respectively). On the other hand, treatment of severe OSAS via CPAP was independently associated with no worse odds of cardiovascular incident than in the healthy men. Doherty *et al.*[100] in a smaller study essentially support these findings by observing a large significant difference in 7.5 year

cardiovascular mortality between a group of treated and untreated OSAS patients (1.9% vs. 14.8%).

1.4 PERFORMANCE AND SAFETY CONSEQUENCES OF SLEEP DISORDERED BREATHING

Sleep Restriction in Normal People

Much of the research investigating human performance and sleep has used the paradigm of total sleep deprivation. However this literature is not particularly useful when looking at the effects of OSAS, where sleep may be shortened and is fragmented, but is not entirely absent.

Experimental sleep restriction is a more applicable paradigm. Belenky and colleagues[102] sleep deprived 66 middle-aged people for 7 days, and then observed them across 3 days of recovery sleep. Participants spent 9, 6, 5, or 3 hours in bed each 24 hour period, and exhibited a dose-response relationship between the level of sleep restriction and poor cognitive functioning. The rate of decline in performance seemed to flatten off somewhat in the 5 and 7 hr groups, after 4 days, which might reflect an adaptation to chronic sleep restriction. This might also help explain the weak relationship between AHI and cognitive impairment in people with SDB/OSAS.

However, on closer inspection, the evidence for a plateau in performance degradation in the less sleep-restricted groups is not as strong as compelling. Furthermore, the considerable influence of inter-individual susceptibility to sleep restriction was not controlled for statistically.

Van Dongen *et al.*, by comparison, have explicitly controlled for the variability between individuals in their performance responses to sleep restriction[103]. In 48 healthy people (21-38 yrs.), 14 consecutive days of sleep restriction to 4, 6 or 8hrs in bed (or three nights of total sleep deprivation) resulted in significant cumulative

decrements in all cognitive tasks. There was a clear dose-response relationship between the severity of sleep restriction and subsequent daytime functioning. Subjective sleepiness ratings, however, did not match the decline in objective performance measures. Van Dongen *et al.* calculate that being awake for more than 15.84 hours results in cognitive deficits that accumulate over experimental days. Optimal sleep duration for maintaining cognitive function is therefore about 8 hours and ten minutes per 24-hour period.

It is important to note that sleep deprivation or restriction may have an even greater effect in people with SDB due to their sleep being more fragmented than that of healthy people.

Sleep Fragmentation in Normal People

Stepanski[104] has recently reviewed 15 studies investigating the effects on daytime functioning of sleep fragmentation in normal sleepers. Sleep fragmentation (usually induced via acoustic tones) induces daytime sleepiness the next day as measured by the Multiple Sleep Latency Test (MSLT), Maintenance of Wakefulness Test (MWT, see Glossary), and single nap sleep onset latencies. Sleep fragmentation also negatively affects cognitive function, mood, and psychomotor performance. Stepanski has interpreted these decrements as being due to the cortical or sub-cortical arousal, and fragmenting of sleep itself. Sleep may need to continuous for a critical period of time to offer recuperative benefits. Some napping suggest that about 10 minutes might be a critical minimum[104, 105]. Alternative explanations as to why sleep fragmentation protocols might cause neurocognitive deficit include the resultant shortening of overall sleep length, or reduction of deeper, theoretically more

recuperative, Slow Wave Sleep and/or REM[106]. Stepansky, however, points out that more recent studies fragment sleep without substantially reducing either sleep length or the relative amounts of SWS or REM[104].

Cognitive Impairment in people with SDB/OSAS

Population Cross-Sectional Studies

Cognitive dysfunction has been studied in the Wisconsin Cohort participants (841 state employed men and woman aged 30-60 years)[107]. Two regression models were constructed, one for psychomotor efficiency, and one for memory. Each controlled for age, gender, and educational status. Psychomotor efficiency was better for people who were college graduates or postgraduates (compared to high school graduates or less), females, younger people, and people with less SDB. This linear model indicates that the association between SDB and psychomotor efficiency follows a dose-response relationship. In the memory model, AHI was not found to predict function independently of the effects of gender, age, and education status. In an inter-group analysis, people in this study with significant SDB (AHI>5) showed worse function in terms of attention and executive function than those without significant SDB, but the size of these effects was not large. The authors report no major differences in terms of memory in the inter-group analysis.

Adams and colleagues report a study investigating 100 people with various levels of SDB with RDI averaging 24 (SD=27.2) recruited from a local population[108]. These researchers identified 4 distinct functional realms, via factor analysis, that summarised the neuropsychological tests used. They named these: declarative memory, signal discrimination, working memory, and set shifting. All models were adjusted for

age. RDI significantly predicted declarative memory (r= -0.22, p=0.03), signal discrimination (r= -0.23, p=0.02), and working memory (r= -0.22, p=0.03), but not set shifting. Time spent below 90% blood oxygen saturation also predicted declarative memory (r= -0.24, p=0.02), and signal discrimination (r= -0.31, p=0.01), but not working memory or set shifting. A model with both RDI and sleep associated hypoxemia was run for signal discrimination and declarative memory. Only the hypoxia index retained significance as a predictor. The authors conclude that the three constructs working memory, declarative memory and signal discrimination are affected in a dose-dependent manner by SDB, after controlling for age, and that these effects are not due to sleepiness (an ESS and MSLT composite). Sleepiness did predict vigilance (r= -.21, p=0.04). Interestingly the polysomnographic Arousal Index was not related to any measure of daytime function, and correlations between daytime function and sleep disordered breathing indices were weak.

Clinic-Based Studies

Kingshott and colleagues investigated the connection between overnight measures routinely calculated as part of polysomnography, and various measures of daytime functioning, in a clinical sample of 150 OSAS patients[109]. The findings of this study support the population findings from the Wisconsin cohort[107]. There were some weak relationships between polysomnographic variables and daytime measures. For instance, there was no significant relationship between AHI, or the arousal index, and daytime sleepiness (MWT or MSLT)[109]. The only nocturnal variable that correlated with daytime sleepiness was the minimum O_2 saturation, which had only a weak relationship with the MWT (r=0.19) and the ESS (r= -0.22).

One of the most useful findings was that the MWT, and not the MSLT, predicts daytime cognitive functioning, albeit weakly[110]. In other words, the relationship between daytime functioning (quality of life, self-reported sleepiness, psychomotor function etc) and sleepiness seems to be more related to the ability to stay awake in a soporific environment, rather than the ability to fall asleep quickly. These relationships are not strong, ranging from r =0.20 to 0.48 (explaining 4-23% of the variance), but all the daytime measures employed were related to the MWT. The MWT is thus a more appropriate test than the MSLT for use in OSAS treatment trials.

Comparisons Made Using Effect Sizes

A major impediment to the review of the evidence concerning cognitive impairment in people with OSAS/SDB is the myriad measures used. No consensus exists as to which daytime functional tests should be used. This lack of standardisation can be serious, as it hampers the preparation of meaningful meta-analyses[111]. A compromise to circumvent these problems is to convert the effects from all measures into effect sizes, which allows comparison of the size of decrements and the effectiveness of treatment between different measures[112, 113]. Studies of OSAS patients compared to controls have shown a wide range of effects in three different general domains; attention, executive function, and memory. Small effect sizes found by Kim *et al.*[107] are contrasted with very large effects in studies such as that by Bedard and colleagues[114]. In general Engleman and colleagues[113] concluded that large decrements were observed in executive function (Effect Size ≈ 0.9) and attention (ES ≈1.0), with moderate decrements in memory-related function (ES ≈ 0.6) in what were largely clinic-based studies.

Engleman and colleagues[113] also reviewed the effects of treatment on the observed decrements. The size of the improvements due to CPAP was small inconsistent across different measures of function. They postulated that cognitive damage caused by OSAS might not be completely reversible with CPAP. Alternatively, the low observed compliance with CPAP may have limited the effect, or the relatively mild OSAS treated in many of the studies they reviewed may have not caused much impairment pre-treatment. These RCTs are reviewed more fully below.

It should be noted that clinical populations generally have more severe daytime symptoms than people identified with the same level of SDB in population studies. Thus, clinical referral bias will tend to lead to overestimation of the effects of SDB on waking function.

Cortical Imaging Studies of OSAS Patients

Kamba *et al.*[115] used proton magnetic resonance spectroscopy of the periventricular white matter to compare cerebral metabolism in normal people to those with OSAS. Twenty-three OSAS patients (11 mild and 12 moderate-severe) had no anatomical abnormalities in their MRI when compared with 15 healthy controls. Patients with OSAS had lower N-acetylaspartate-to-choline ratios than controls, and patients with moderate-severe OSAS were worse than those with mild OSAS. These altered cerebral metabolite ratios indicate the presence of some sort of cerebral damage, which the authors hypothesise is due to SDB-induced repetitive hypoxia. The functional significance of these differences in metabolite ratios remains unclear.

Bartlett and colleagues[116] have compared the left hypocampii of 8 adult males with OSAS to 5 age matched controls, by using high resolution magnetic resonance spectroscopy (proton MRI). OSAS patients (AHI>15, min SaO_2 <90%) had lower levels of creatine-containing compounds, and the level of these compounds were correlated with the severity of OSAS and observed neurocognitive decrements. It was argued that the levels of these metabolites are thus, functionally and clinically, significant markers of disease severity. There were, however, large variations in the N-acetyl-containing-compounds (markers of neuronal integrity) within the OSAS group, perhaps indicating individual susceptibility to the effects of OSAS.

Macey and colleagues[117] compared brain morphology in 21 male patients with largely moderate to severe OSAS (Mean RDI=38 SD=24) with 21 control subjects, by using high resolution T1 weighted magnetic resonance imaging. Grey matter was reduced in OSAS patients in the frontal and parietal cortex, the temporal lobe, cerebellum, hippocampus and anterior cingulate. These areas are not only implicated in the cognitive deficits observed in people with OSAS, but also in the fine motor regulation of the upper airway. These observed differences may thus be both a cause of and/or a result of SDB[118]. The amount of damage observed in the OSAS patients was in a rough dose-response relationship with the severity of the SDB.

The authors[117] observed that some of the differences observed are bilateral and diffuse in nature, as would be expected from repetitive hypoxic insult. However some of the other changes are not the sort one might expect from the hypoxic action expected of OSAS over time. Specific sites of damage on one side of the cortex, such as in the left ventral lateral frontal cortex (including but not exclusive to Broca's area), are indicated as a possible cause of OSAS, as this area is known to help control upper

airway motor function and speech production. Damage observed in the cerebellum is also unilateral, in this study. Damage to the cerebellum has been known to induce OSAS, which is not surprising given it's role in fine motor control during wakefulness. Supporting this concept was the unusually large proportion of the people in Macey *et al.* sample who reported childhood speech problems.

Accident Risk in People with SDB/OSAS

Motor Vehicle Accidents

Given that cognitive impairment, particularly in vigilance-related tasks, is observed in people with OSAS, a number of research groups have looked at whether OSAS increases the risk of motor vehicle accidents (MVAs). The peak in OSAS prevalence in the community seems to be in middle age which is, coincidentally, where the lowest risk for MVAs occurs. Furthermore, some clinical patients may have been referred for specialist attention only after having fallen asleep at the wheel and crashed, which may inflate the estimation of risk, and later inflate the apparent effectiveness of treatment in reducing risk. Separating out the effects of sleepiness in addition to SDB is usually not possible in clinical case control studies, which make up the bulk of the evidence that OSAS increases the risk of MVA.

The best available evidence currently comes from the Wisconsin cohort. Male Wisconsin state workers with an AHI>15 had an odds ratio of 7.3 times the risk for multiple automobile accidents (3.4 times the risk of at least one accident) when compared to workers without SDB[119]. However, the increased risk did not follow a clear dose-response relationship, and was not found in women. Possibly this is due to

reduced statistical power with fewer women in the cohort and that women are less likely to have SDB, and also less likely to have MVAs.

Importantly in this study, multiple measures of sleepiness did not independently increase the risk of MVA after controlling for SDB. Given that sleepiness is the putative causal mechanism between SDB and MVA, either the measures of sleepiness failed to capture real-world risk, or there is a mismatch between actual sleepiness and the subjective experience of sleepiness in this community-based sample. Sleepiness might be reasonably expected to be situation specific[120]. Thus asking about sleepiness in a motor vehicle might be more specific than asking about general sleepiness. Our own population based studies indicated that the two questions in the Epworth Scale (Question 4 and 8) that address sleepiness in a motor vehicle predict MVA involvement, whereas the whole Epworth Scale does not[121, 122].

Other evidence supporting the hypothesis that OSAS increases the risk of MVAs is reviewed elsewhere (see Appendix 3[123]) and generally indicates that people with severe OSAS (conservatively estimated) have about 2-3 times the risk of MVA compared to people of the same age and gender who do not have OSAS. However, there is little evidence of a clear dose-response relationship between worsening levels of SDB/OSAS and increasing risk of MVA, except from a single letter reporting MVA risk in a clinical sample[124].

Clinical studies indicate that the treatment with CPAP reduces crash risk to control levels[125, 126]. These are not randomised controlled studies, but they offer the best evidence available (or likely to be available given ethical constraints) that CPAP treatment reduces the risk of MVA in people with severe OSAS. These studies are

perhaps still subject to regression to the mean effect, and I have argued that the published odds ratios are probably inflated because of this[123].

Industrial Accidents

In a 10-year follow-up of industrial accidents in a group of 2874 working men (aged 30-64 in 1984) from Uppsala in Sweden, Lindberg and colleagues[127] reported a 73.8% response rate in 1994. Industrial accidents were injury causing events that were recorded in the Swedish national registry to which reporting is mandatory, although not complete. The aim of this study was to assess whether sleepy snorers, who presumably have a high risk for having OSAS, have higher rates of occupational accident and injury. Men were classed as sleepy if they answered within 3-5 on a 5-point scale to the question "How often do you fall asleep involuntarily for a short period during the day, for example when there is a pause at work?". Snorers were those who answered within 3- 5 on a 5-point scale to a question about how often they snored loudly and disturbingly. Men who were sleepy snorers were 2.2 times more likely to be involved in an industrial accident over the ten year period (95% CI =1.3-3.8), after controlling for the effects of age, years at work, body mass, smoking, alcohol dependence, blue collar vs. white collar work, organic solvent exposure, noise exposure, exhaust fume exposure, whole body vibration, and shiftwork. No association was found between industrial accidents and either snoring or sleepiness alone. Again this is somewhat puzzling, given that sleepiness is thought to be the causal mechanism, but it is consistent with the findings from the Wisconsin cohort[119]. The overall findings of Linberg and colleagues' study need to be interpreted with caution as the lack of polysomnographic measurements makes it impossible to rule out significant misclassification of OSAS status.

Poor Correlation Between SDB Severity and Functioning.

A number of factors might contribute to the lack of consistent dose-response relationships between SDB and daytime dysfunction.

- Clinical Referral Bias. Only those significantly impaired make it to a sleep clinic.
- Current measures of SDB, such as the AHI, are not an accurate correlate of the damaging aspect/s of SDB.
- Current measures of daytime function are inadequate.
- There is a large amount of individual variability in the effects of SDB on individuals.

1.5 TREATMENTS FOR OBSTRUCTIVE SLEEP APNOEA SYNDROME

The gold standard for the treatment of OSAS is Continuous Positive Airway Pressure (CPAP). As this treatment is the focus of this thesis it will be discussed last after first describing the alternative treatment modalities.

Oral Appliances and Mandibular Advancement Splints

Oral Appliances (OAs) are worn during sleep and are usually individually moulded to fit the teeth. The goal of these devices is to reduce SDB by moving the mandible forward in relation to the maxilla, enlarging larger upper airway which is thus less prone to collapse. The degree of advancement is individualised according to what the patient can tolerate, offset against the potential for occlusal changes[128] and discomfort especially to the temporomandibular joint. However, general consensus is that the greater the forward movement, the greater the reduction in SDB[129].

The effectiveness of Oral Appliances has been subject to a Cochrane review[130]. In general it seems that well-designed OAs about halve the AHI in patients with moderate-to-severe OSAS. Halving the AHI for the whole night might be preferable to CPAP eliminating the AHI for smaller proportions of the night for patients unable to fully comply with CPAP therapy[130]. The major benefit of OAs is that they are generally better tolerated than CPAP, have fewer side effects, and are less visually and aurally intrusive for some patients[129, 131, 132]. Some recent evidence suggests that these devices reduce blood pressure during sleep[89, 90]. This is extremely promising for the treatment of mild-moderate OSAS, where considerable uncertainty exists as to the optimum treatment strategy and any back-up strategy. The variety of

OAs available makes it difficult to assess their overall effectiveness at reducing SDB except on a device-by-device basis.

Surgical Methods

Surgical methods used to alleviate OSAS by tissue ablation are all attempting to increase airway calibre or to remove airway obstruction caused by soft tissue abnormality. However, the most effective surgical method for treating upper airway collapse during sleep is a tracheostomy. Tracheostomies completely bypass the upper airway, thus eliminating obstructive events during sleep. They do however require constant vigilance to maintain patency and are unsightly. Tracheostomy was a more common treatment modality during the 70's and early 80's in patients with very severe OSAS, but has fallen out of favour as CPAP therapy became more widespread[133]. It remains a surgical option in patients with severe OSAS and cardiovascular comorbidities who are unable to tolerate CPAP[133, 134].

Other surgical methods might be particularly effective in patients with severe OSAS, who are not obese, and whose airway obstruction is largely caused by an anatomical abnormality[135]. Anatomical features that are often amenable to surgical removal or minimisation include large tonsils and/or adenoids (adenotonsillectomy), large uvula and/or floppy soft palate (uvulopalatopharyngoplasty-UPPP or U3P), and mandibular and maxillary insufficiency (Maxillomandibular advancement osteotomy). Upper airway surgical techniques might be of particular relevance in Asian populations, where obesity is not a prime determinant of sleep disordered breathing, but might be more closely related to maxillary insufficiency. However, as is common with many surgical techniques, the various methods have not largely been tested in RCTs. A 2003

Cochrane review found no randomised or quasi-randomised controlled trials investigating the effectiveness of any upper airway surgery in treating Obstructive Sleep Apnoea Syndrome[136].

Some surgical methods aimed at reducing morbid obesity, as a beneficial side-effect also reduce SDB, which is often observed in these patients[137]. However, this approach is not relevant to the majority of patients who are slightly overweight or obese and have mild-moderate SDB.

Conservative Management and Sleep Hygiene

Conservative advice is non-invasive lifestyle modification generally aimed at reducing the impact and severity of OSAS[138]. It typically includes avoidance of the supine sleep position, avoidance of hypnotics and sedatives (including alcohol), consistant bedtime routine, avoidance of stimulants (including caffeine), and weight loss where appropriate. The advice given as part of conservative measures is generally sensible extrapolation from good physiological data. It is, however, far from clear whether people are able to comply with this advice and whether compliance produces any benefits for cardiovascular health or quality of life. A Cochrane review recommended that RCTs are required to show whether these types of advice are separately or conjointly useful for treating OSAS/SDB[139].

Continuous Positive Airway Pressure (CPAP)

Continuous Positive Airway Pressure (CPAP) is a small pump device, attached to a hose and mask interface, that seals over the nose (or mouth or both) and delivers pressure via hosing and a mask interface. By constant application of positive pressure to

the upper airway, CPAP offsets the pathological negative pressure imbalance that immediately precedes airway collapse in people with SDB. CPAP was developed in the early 80's[140] and has in essence remained unchanged since then, despite a race to in improve the device. This technological race has spawned numerous new interface mask designs, each advertised as being more comfortable than the last. It has also spawned new CPAP-like modalities including, BiLevel (Or BiPAP, which is PAP with a different pressure at inhalation and exhalation), AutoPAP (self titrates the therapeutic pressure up and down throughout the night), and C-Flex (pressure modulation at exhalation- see Chapter 4 for comparisons of the devices). When worn correctly, and properly titrated to an individual's pressure requirement, CPAP is nearly 100% effective at abolishing apnoeas and hyponoeas and the resultant intermittent hypoxia[140]. However a number of patients have comfort issues with the device and compliance is often sub-optimal.

1.6 THE EFFECTIVENESS OF CONTINUOUS POSITIVE AIRWAY PRESSURE.

The Wright *et al.* Meta-Analysis and it's Consequences.

In 1997 Wright *et al.* published a systematic review and meta-analysis to ascertain what role if any OSAS might have in important health related outcomes and, to ascertain the effectiveness of CPAP in treating the syndrome[12]. They found poor evidence that OSAS was related to any vascular outcomes or mortality, due to failure to control for obesity and smoking in the studies examined. Many studies diagnosed OSAS after a comorbid condition had already been identified. The better-designed studies failed to find associations with objective detrimental outcomes. Weak, but consistent, evidence linked OSAS with poor simulated driving skills and real-world motor vehicle crashes, but only among people with severe OSAS. The entire research literature was found to be troubled by the confounding effects of common cardiovascular risk factors that have tended to be found in the so called "Syndrome X" cluster; systemic hypertension, insulin resistance, hyperlipidemia, and central obesity. The authors lamented the lack of longitudinal cohort data investigating OSAS, but should have been aware that such data were forthcoming[3, 141].

With regard to whether OSAS could be successfully treated with CPAP, Wright *et al.*[12] concluded that there was no real evidence of efficacy, other than for the most severe levels. This conclusion was at odds with the only RCT available at the time[142].

As a direct result of these findings, public funding for CPAP treatment was halted in the United Kingdom, despite the protests of those in the sleep medicine community in the UK and abroad[141, 143-147].

An Australasian meta-analysis[148] came to quite different conclusions about the effectiveness of CPAP in dealing with the same lack of evidence. However it used unusual methodology and even the final version[111] included recommendations for treatment that did not match the evidence identified in the systematic review.

Since these two systematic reviews, a number of well-designed trials have been published, including 6 RCTs dealing exclusively with mild or mild-moderate OSAS. The seventh such trial is reported in Chapter 2 and Appendix 3[38], and uses sham CPAP as a control, as recommended by Wright *et al.*[12].

Randomised Controlled Trials of CPAP Effectiveness for Treating Severe OSAS

Effects on Sleepiness

The best evidence that CPAP reduces sleepiness in people with OSAS comes from a recent systematic review and meta-analysis of all placebo and conservatively controlled RCTs[149]. A systematic search identified 12 RCTs that included patients with a wide range of OSAS severity. Sleepiness was measured using either the Epworth Sleepiness Scale (n=11), the Multiple Sleep Latency Test (n=6) or the Maintenance of Wakefulness Test (n=2). Meta-analysis indicated that scores from the Epworth Sleepiness Scale improved by a mean of 2.94 points on CPAP, after adjustment for placebo effects (95% CI 1.61-4.26, $p<0.001$). A pooled metric, based on the mean improvements in latencies of either the MWT or the MSLT (objective sleepiness), showed an improvement of 0.93 minutes (95% CI 0.10-1.76, $p=0.04$) with CPAP. After excluding RCT's, with exclusively mild-moderate OSAS (AHI 5-30/hr), the ESS improved by a mean 4.75 points (95% CI 2.97-6.53, $p<0.001$). The combined MSLT/MWT metric would probably have had a larger effect if the three studies with

severe OSAS had been separately analysed and presented[142, 150, 151]. The reviewers concluded that CPAP was an effective treatment for both subjective and objective sleepiness in people with severe OSAS, but that they were unable to confirm that this benefit extended to those with mild-moderate OSAS. Since this most recent meta-analysis, no further randomised trials focussing on severe OSAS patients have been published.

Effects on Blood Pressure

Six trials have looked at the effectiveness of CPAP for reducing blood pressure in people with, usually, severe OSAS. A further instructive non-randomised study was discussed earlier[58].

The first RCT used a three-week crossover design to investigate the effects of CPAP compared to oral pill placebo, on ambulatory blood pressure in 13 severe OSAS patients[81]. There was no effect on blood pressure due to CPAP. However in the sub-group of patients (n=5) who did not have the customary overnight dip in blood pressure, daytime BP improved more on CPAP than on placebo (4mmHg mean difference, p=0.01). The authors interpreted these findings to indicate that CPAP would not reduce BP in a heterogenous sample of people with SDB, but might do so in those most at risk of cardiovascular disease.

Dimsdale *et al.*[82] attempted to compare the effectiveness of CPAP versus sham CPAP in reducing BP. Unfortunately, the attempted placebo CPAP device actually reduced the indices of sleep disordered breathing and was in reality an under-titrated CPAP ($2cmH_2O$). However, full therapeutic CPAP reduced mean arterial

pressure by about 4mmHg after 7 days of treatment. The partial sham CPAP device also reduced mean arterial pressure, but only by about 2mmHg.

Similar findings have been reported by Becker and colleagues[78], using a similar partially therapeutic CPAP device compared to fully titrated CPAP. Ten mmHg reductions in mean 24-hour BP were seen in the CPAP arm, in both systolic and diastolic pressure. In contrast patients on partially therapeutic CPAP showed no reductions in BP.

Faccenda and colleagues[80] have reported the effects on blood pressure of CPAP, compared to a pill placebo, in a 4 week crossover trial (n=68, AHI median 35). All patients were normotensive at baseline. Mean 24-hour diastolic pressure dropped by 1.5mmHg more on CPAP than placebo (p=0.04). Mean systolic pressure dropped by a similar amount, but this was not significant (1.3mmHg, p=0.19). Those who used CPAP more than 3.5 hours per night on average, and might therefore be expected to benefit more, had an average net drop in diastolic mean pressure of 3.7 mmHg (p=0.03). The magnitude of the effect on systolic and mean pressure was similar, but not statistically significant. The group with the most severe oxygen desaturations (i.e. >20/hr of at least 4%) benefited significantly on all BP measures with reductions of 4mmHg (systolic), 5mmHg (diastolic) and 3.4 mmHg (mean, all $p<0.05$).

Barbe and colleagues[83] tested CPAP against sham CPAP in a group of 55 patients with severe SDB (AHI> 30, mean 56/hr) but no daytime sleepiness (Epworth Sleepiness Scale <11). After 6 weeks of treatment in a parallel design, they found no differences in 24-hour BP (neither mean, systolic, or diastolic), nor in a daytime or

night-time comparison. Patients were, however, relatively normotensive at baseline (diurnal mean 130/82mmHg).

The sixth RCT was a sham CPAP controlled parallel study of the effects on ambulatory BP of CPAP in 118 men with mostly severe OSAS, diagnosed by >4% blood oxygen desaturations[79]. One month of CPAP therapy reduced mean arterial blood pressure by a significant 2.5 mmHg (3.3mmHg taking into account s slight worsening on placebo, p=0.0013). The benefits were seen in both systolic and diastolic BP. CPAP was more effective in people with severe OSAS (>33 4% dips/hr). It was also more effective for those who took hypertensive medication at baseline (6.6 mmHg, n=22), although was seemingly independent of the baseline BP values. High compliers with CPAP (>5hrs/night) experienced a significant fall in mean BP (4.9 mmHg, p=0.001). The authors conclude that typical OSAS patients receiving treatment in the UK can significantly reduce blood pressure by using CPAP.

Effects on Neurocognitive Function

General neuro-cognitive testing in RCTs has been fraught with problems of measurement inconsistency. This issue can be seen even in very similar tests such as the MSLT and MWT. These tests use almost identical protocols, and purport to measure a similar ability, but have a surprisingly poor correlation of (at best) r=0.43[110], and are differentially sensitive to CPAP treatment (see Chapter 3 and also Sangal *et al.*[152, 153]). Many cognitive tests are also prone to practice effects. A notable exception are simple tests of reaction time, such as the Psychomotor Vigilance Task (PVT)[154].

The first RCT[142] employed the National Adult Reading Test (NART), the Wechsler Adult Intelligence Scale (WAIS-R; digit symbol substitution and block design

tests), trail-making A and B, Rapid Visual Information Processing Test (RVIPT), Steer Clear, the Paced Auditory Serial Addition Task (PASAT at 2 and 4 second intervals), and eight choice reaction time. Further tests for Verbal fluency (Borkowski test) and memory (visual retention Benton revised test) were also employed. Comparing 4 weeks of treatment with CPAP to 4 weeks of treatment with placebo, in a crossover trial, it was found that trail-making B, Digit symbol substitution, and IQ decrement improved more on CPAP. Despite a 5-hr practice session, learning effects were observed over the trial in Steer Clear, trail-making B, RVIPT, digit symbol substitution, and the 2-second PASAT. Randomisation of treatment order controlled for these effects. The authors interpreted these findings as an indication that cognitive deficits in vigilance, mental flexibility, general intelligence, and coding speed were all reversible, at least partly. With longer treatment duration and/or higher compliance with CPAP, the remaining deficits might have been fully reversible.

A second trial by the same group[151] investigated 23 patients with severe OSAS in a RCT crossover trial comparing CPAP to a pill placebo. This study used 10 cognitive outcomes measures, including Steer Clear, Trail-making B, digit symbol substitution, Block design, performance IQ decrement, Benton visual retention test, verbal fluency, PASAT (2 secs), and Rapid Visual Information Processing. None of these tests showed significant improvement on CPAP compared to placebo.

Barbe and colleagues, tested the efficacy of CPAP in people with severe SDB but without daytime sleepiness[83]. They employed Steer Clear, Wechsler adult intelligence scale (digit symbol substitution test, block design, and digit span), PASAT (1, 2, 3, 4-second versions), Wechsler Memory (mental control and verbal paired

associated), and trail-making (A and B). After 6 weeks of treatment only the PASAT (2-second interval) showed significant improvement on CPAP compared to sham CPAP.

Henke *et al.*[155] compared CPAP to sham-CPAP in a partial crossover trial with 46 people with severe OSAS (mean AHI=65). None of the following measures improved with a mean 16 days of CPAP therapy: trailmaking A and B, digit symbol substitution, digit span backwards and forwards, Steer Clear, Controlled oral word association, Medical College of Georgia complex figure recall, and selective reminding test (total words recalled).

Hack and colleagues[156] measured secondary reaction time whilst using a driving simulator, in patients with severe OSAS on CPAP or sham CPAP. Secondary reaction time improved from 2.8 secs to 2.3 secs during CPAP treatment. This improvement approach statistical significance when compared to the 0.1 second improvement during placebo treatment (p for the difference=0.07).

Bardwell and colleagues[157] found that only 1 of 22 neuropsychological tests they employed improved more on CPAP than on sub-optimal CPAP, after 1 week of treatment (Digit vigilance time). This difference might have been due to chance alone, but a partially therapeutic 'placebo' device ($2cmH_2O$) might have biased in favour of false negative findings.

In summary, a multitude of neurocognitive measures have been employed in various trials of CPAP in severe OSAS patients. It is far from clear whether these measurements capture functionally relevant traits, or are especially susceptible to the effects of OSAS, or are reversible by OSAS treatments. Engleman and colleagues[113]

have meta-analysed their own 3 RCTs, which have all used the same measures, and found reversibility in most outcomes. However, treatment did not restore functioning to normal levels. This might be due to the impairment not being fully reversible, or to the generally low levels of compliance with CPAP therapy in these RCTs (ca. 2.9 hours of effective pressure per night), or to the treatment periods being too short for full recovery.

Effects on Simulated Driving Ability

Hack and colleagues have reported a randomised parallel trial of CPAP compared to sham CPAP in 59 patients with severe OSAS comparing simulated driving performance[156]. Variability in steering position improved more on CPAP than placebo ($p=0.03$). Deterioration in steering performance over the test also improved more on CPAP than on placebo ($p=0.04$). However the number of off-road events (i.e. crashes) was not different between groups. Epworth Scores and a substantially modified non-EEG version of the Maintenance of Wakefulness Test (the OSLER[158]) were also markedly improved on CPAP, after adjustment for placebo.

The Steer Clear computer program is a very simple driving simulator which has been used in three RCTs of CPAP in people with predominantly severe OSAS[83, 142, 155]. Engleman and colleagues[142] found that the number of obstacles hit was lower in the CPAP group than the placebo group ($p=0.01$). Barbe and colleagues[83] found no improvements when CPAP was compared to sham CPAP for patients with AHI>30 and no significant daytime sleepiness. Henke and colleagues[155] also found no difference comparing CPAP to sham CPAP for patients with severe OSAS.

Results from individual patients need to be interpreted with care, as driving simulators are prone to learning effects and can massively overestimate the accident risk posed by females[159], who are at much lower risk of real world crashes. Realistic driving simulators currently being developed for sleep research and are less prone to this problem[160].

Effects on Quality of Life and Mood

A number of trials exist have attempted to measure the effects of CPAP on mood and quality of life in people with a wide range of OSAS severity.

Engleman and colleagues[151] compared CPAP to an oral placebo in a four week crossover trial with 23 patients who had moderate to severe OSAS (AHI>15, Mean 43±37/hr). After controlling for placebo effects, none of the mood measurements improved (Hospital Anxiety and Depression Scale- Anxiety and Depression, General Health Questionnaire-28, or the Nottingham Health Profile- Part 2).

Ballester and colleagues[161] compared the effects of CPAP and conservative therapy to conservative therapy alone in a parallel randomised study with 105 patients who had Severe OSAS (Mean AHI=56). The Nottingham Health Profile improved in two of six subscales, Energy ($p<0.005$) and Social Isolation ($p=0.03$).

Jenkinson and colleagues[150] compared CPAP to sham CPAP in 101 patients with moderate to severe OSAS (ODI 4% > 10/hr, mean 31/hr). The SF-36 was susceptible to placebo effects, but scores on four of eight subscales improved more on CPAP than placebo (Physical Role, Mental Role, Mental Health, and Vitality/Energy,

all $p<0.01$). Both the Physical and Mental Component summary scales also improved on CPAP more than placebo (both $p<0.01$).

Yu and colleagues[162], using the same sub-therapeutic CPAP control reported by Dimsdale and colleagues[82], found that one week of CPAP therapy significantly improved mood in 5 of 6 subscales in the Profiles of Moods States (POMS) and also the overall Mood Disturbance (total POMS). However, these improvements were not greater than those observed in the sub-therapeutic CPAP arm.

In another sub-optimally controlled study by the same group[163], CPAP was no more effective than sub-therapeutic CPAP at improving health quality of life as measured by the Medical Outcomes Survey (full version of the SF-36). Significant 'placebo' effects were observed, but may in fact have been elicited by the physical action of the placebo in reducing sleep disordered breathing. It is difficult to conclude anything about the efficacy of CPAP treatment from any of these studies, due to the poor design of the placebo treatment.

Barbe and colleagues[83] have reported the findings of a multi-centre parallel trial of CPAP (vs. sham CPAP) in 55 patients with severe asymptomatic SDB (mean Epworth score 7.0 at baseline, mean AHI=55). After 6 weeks of therapy, no differences were observed between the two groups according to the full Functional Outcomes of Sleepiness Questionnaire (FOSQ) or the Physical and Mental Component summary scores from the SF-36.

Montserrat and colleagues[164] used a partial crossover design to compare CPAP to sham CPAP in 45 patients with severe OSAS (Mean AHI 54/hr). Parallel

group results from this trial showed that 2 of 4 subscales from the FOSQ (General Productivity and Vigilance) improved more on CPAP than on placebo, but that no measure from the SF-36 improved, although the Physical functioning subscale approached significance ($p=0.057$). There were notable placebo effects in all FOSQ scales except General Productivity (all $p<0.06$). Four of 8 subscales (Role Emotional, Social Functioning, Bodily Pain, and Vitality; all $p<0.06$) in the SF-36 and the Mental Component summary scale were also sensitive to placebo effects ($p<0.01$).

Faccenda *et al.*[80] also reported FOSQ improvements in their study (see above, under blood pressure). Scores on the Total FOSQ, Vigilance, Activity Level and Social Outcomes were higher after treatment in the CPAP group than the placebo group (all $p<0.03$).

It seems that CPAP may improve quality of life (SF-36 and FOSQ in particular), in people with severe OSAS. However, the sub-scales that improved, did not seem to be stable from study to study. Effects on mood scales seem to be relatively heterogenous, and prone to placebo effects.

Qualities of a Good Clinical Trial of CPAP

Randomisation to Treatment

Randomisation removes the potential for systematic bias by making the likelihood of being placed in any treatment group the result of chance alone. Randomisation only works when the allocation sequence is hidden from those enrolling patients in a trial[165-167].

Hidden Treatment Sequence

In studies where the treatment sequence can be deduced, patient groups can be deliberately, or unconsciously, altered to favour one or another treatment group. In CPAP research, false results in favour of CPAP might be engineered via placing patients, who are thought to be likely to have higher compliance, onto the active treatment, or by preferentially placing particularly sleepy patients on CPAP.

Choice of Control Comparison

The choice of a placebo has been the subject of some controversy[80, 168-170]. The Edinburgh group have favoured orally ingested placebos (pills) as they have a neutral effect on sleep[80, 151, 170]. Two Australian studies have taken a similar stance[90, 171]. Other commentators have called for sham CPAP placebos[12, 168], where CPAP machines are tampered with as to deliver a pressure so low that it does not reverse airway obstruction during sleep[169]. It is argued that the sham device more accurately controls for the negative effects of wearing a mask, and for the considerable potential placebo effects associated with a complex behavioural intervention, such as CPAP [12, 168].

A further benefit of sham CPAP that has not been mentioned by other commentators, is that the placebo effects of different levels of compliance can be separated from the actual physiological effects of different levels of compliance with CPAP. Patients who use a treatment expect some benefit. Those who do not use a treatment expect no benefit. Thus, the use of sham CPAP is a means for examining the dose-response relationship between this psychological expectation, and real benefits.

Placebo controls are still generally required in any trial of the effectiveness of CPAP treatment. Trials are no longer required to show that sleepiness is reduced in

those with severe OSAS[149]. Effectiveness of treatments on other outcomes of interest, such as the retrieval of cognitive function, quality of life, or cardiovascular health, should still be examined in placebo controlled studies, as it is not yet clear what the effects of CPAP are on these outcomes.

Double Blinding

Double blinding controls not only for the expectations of patients, but also for the potentially more accurate preconceptions/preferences/biases of investigators, that may affect the measurement of outcomes. In CPAP research, the double blinding principle is particularly hard to enforce. The need to actively and individually titrate CPAP means that sleep technologists (those responsible for CPAP titration) interact directly with patients and have direct knowledge of the efficacy of the treatment being used. Furthermore, sleep physicians will also have direct knowledge of the efficacy of a device (CPAP will eliminate apnoeas and hypopnoeas, sham CPAP will not), and will also likely have some interaction with the patient. Pragmatically, the best double blinding that can be achieved in a CPAP trial is to maintain the blinding of those who collect the major outcome variables, particularly subjective outcomes such as quality of life, mood, and sleepiness measures. Although true double blinding is extremely difficult in a CPAP trial, the maintenance of single blinding (patient blinding to treatment effectiveness) is possible.

1.7 GENERAL SUMMARY OF LITERATURE

The literature post Wright *et al.*[12] has greatly clarified many issued raised in that review. Cross-sectional studies decrements in a wide range of outcome measures, even with SDB of mild-moderate severity. Longitudinal cohort data indicates that SDB is associated, in a dose-dependent manner, with the risks of developing hypertension. Coupled with other lines of evidence, it therefore seems highly likely that SDB causes cardiovascular disease including heart failure, stroke and myocardial infarction.

CPAP is a generally effective treatment for severe OSAS (AHI>30/hr) but the improvements observed in important outcomes do not restore function to normal levels. Suboptimal compliance is possibly a key factor. The decrements might be permanent in nature and not be reversible. In this case, CPAP could be used more as a prophylactic treatment for cognitive decline, or cardiovascular disease in particular. However, there are no data on the effectiveness of this approach. Further problems remain, there is little consistent RCT based data regarding the effectiveness of CPAP, or other treatments, for a larger group of potential patients with mild-moderate OSAS (AHI 5-30/hr). An even larger proportion of the population have various degrees of SDB, without sleepiness, but even CPAP seems ineffective in even the severe members of this group[83]. There is little effort to systematically identify this larger asymptomatic group, which is also at higher risk for cardiovascular diseases.

1.8 AIMS OF THE THESIS

This thesis reports the findings from 2 clinical trials and a systematic review and meta-analysis. These studies target two concepts neglected in sleep medicine; clinical and technological drift.

In this thesis clinical drift is defined as "the practice of applying treatment demonstrated to be effective and safe to a different group of patients for whom the treatment has not been demonstrated to be effective or safe". The effectiveness of treatment can vary between diseases/syndromes, and within diseases/syndromes due to factors associated with age, gender, ethnicity, or severity. Typically, this practice has been defended through the allowance of "off-label" prescribing, where a disease is treated with a device or drug which is not officially indicated for that purpose, but which the clinician believes is in the best interests of the patient[172]. This is ethically defensible where no randomised controlled trials exist to guide clinicians, especially where a patient is in a very small class of people with complex co-morbid conditions. This practice is orthodox and is not usually experimental or risky[172]. Off-label use usually means that the use of the treatment for the combination of disease/syndrome has not yet been officially sanctioned, usually due to lack of information. Ideally, this practice should not extend to entire classes of patients such as the elderly, paediatric, women, teenagers, or those with various severities of syndrome. These groups warrant specific investigation before the widespread adoption of a particular treatment option in that class. The recent controversy surrounding the Selective Serotonin Reuptake Inhibitor class of pharmaceutical treatments (SSRIs), and the potentially increased risk of suicide or suicidal thoughts amongst teenagers, is a good example[173, 174]. That situation was further complicated by the data having been gathered, but selectively published, by manufacturers[175].

Sleep medicine probably has a somewhat different problem because much needed information has never been gathered. During the 1980s CPAP was demonstrated to be an effective treatment for the very severe OSAS cases reaching sleep clinics at that time[140]. Use of the treatment spread rapidly and to different patient groups. The first placebo controlled RCT was not published until 1994[142], and included a full range of AHI severity (Apnoea Hypopnoea Index 7-129/hr). Treatment guidelines from around the world[111, 176, 177] have recommended CPAP treatment despite the uncertainty about the size of the benefits of CPAP for patients with less severe OSAS[12]. The guidelines blanket all adult patient classes despite the lack of trials in important patient sub-classes. The lack of demonstrated benefit to those with mild-moderate OSAS is also problematic, with guidelines generally recommending CPAP treatment in this class. There are now 7 published randomised controlled trials of CPAP exclusively in patients with mild or mild-to-moderate OSAS. One of these constitutes the first part of this thesis, and is the first sham CPAP controlled crossover trial of CPAP in any severity of OSAS[38]. The second part of the thesis combines all seven studies in a meta-analysis based on a systematic review, in order to quantify the effectiveness of CPAP for treating daytime sleepiness in patients with mild-moderate OSAS.

The other uniting theme, in the research reported in this thesis, is that of technological drift. The development of new technology can seem to be random and/or unplanned or unfocused toward human wants or needs- hence the idea of an aimless drift in technological advancement. In sleep medicine, and in healthcare in general, it manifests as the problem of whether modifications to existing treatments provide cost-effective benefits to patients.

This concept is illustrated in the marketing of healthcare screening programmes. In the United States, new technological advancements such as Magnetic Resonance Imaging (MRI) allow full body scans for sub-clinical abnormalities. Prostate Specific Antigen (PSA) tests seem to be in a similar category, in widespread use, but of dubious net worth. These techniques are costly and it is not certain that they save lives, by comparison with waiting until the abnormality manifests clinically in the traditional manner. In the meantime, these techniques produce a number of false positive cases causing significant worry and emotional distress to patients. In some of these screening industries, the original wait and see approach might be just as effective, or more effective and considerably less costly, than the new high tech approach.

In sleep medicine the same problems exist. CPAP is a very effective device for the suppression of sleep disordered breathing, virtually eliminating the nocturnal manifestations of the condition. But CPAP is only fully effective when it is worn. Sub-optimal compliance with CPAP has lead to a technology race between the main CPAP manufacturers to develop ever newer and more expensive mask interfaces and CPAP-like devices, on the basis that they increase compliance with treatment. Heated humidification of air in the CPAP circuit is one such advance that is marginally, but significantly, beneficial for increasing compliance[39, 178]. Other additions with less certain benefits to compliance which are nevertheless used to increase compliance are pressure ramping, automatically titrating CPAPs (AutoPAP), and Bilevel PAP machines for uncomplicated OSAS.

C-Flex™ is a new type of CPAP-like device designed to prevent airway occlusion during sleep, whilst lowering exhalation pressure (Respironics Inc, NC, USA). C-Flex is the subject of the third study reported in this thesis. This new device is

advertised by the manufacturer as increasing compliance by one hour per night compared to CPAP, but no published evidence exists to support this claim. The device is already in widespread use, is probably no more technically effective in reducing the nocturnal manifestations of sleep disordered breathing than CPAP, and is approximately 50% more expensive. An interrelated problem is that the additional functional benefit of an extra hour of CPAP per night is unknown.

Improving Evidence Based Sleep Medicine

> *Evidence based medicine is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research*
> ***Sackett et al. BMJ 1996 v.312 pp 71-72[179]***

EBM advocates generally list meta analyses based on systematic reviews of randomised controlled trials (Chapter 3), and randomised controlled trials (Chapter 2 and 4), as being at the top of the evidence hierarchy when evaluating the benefits of a treatment (i.e. levels 1a and 1b). Cohort studies, cross-sectional studies, case-control studies, case series and individual cases are progressively less reliable treatment evidence and should generally not be relied upon when RCTs are available. This hierarchy seems to be supported by the overall citation pattern in leading health journals where meta-analyses and then RCTs are the most commonly cited study designs[180].

During the 1950's, English physician Archie Cochrane demonstrated in the streets of London with a sandwich board demanding free and effective healthcare for

all. It took some time for Cochrane's views to be widely adopted. He is now often given credit as one of the pioneers of EBM and the international collaboration for systematic reviews and meta-analyses is named for him (www.cochrane.org). The concepts of free and effective healthcare are a powerful combination of responsibilities, especially in societies where healthcare is mostly publicly subsidised and centrally rationed (such as New Zealand, Australia and the United Kingdom). Effective, rather than ineffective, healthcare then becomes an ethical imperative. To continue to use healthcare techniques that are unproven (or worse, demonstrated not to work) is, at best, to deny effective healthcare to people in need and to replace true care with placebo effects. At worst it is to subject people to damaging, toxic, or fatal side-effects, without objective benefit. EBM is very simply the sifting of all medical and para-medical practices into those which work and those which do not.

> *"scientific medicine is **defined** as the set of practices which submit themselves to the ordeal of being **tested**. Alternative medicine is defined as that set of practices which cannot be tested, refuse to be tested or consistently fail tests. If a healing technique is demonstrated to have curative properties in properly controlled double blind trials, it ceases to be alternative. It simply ...becomes medicine"*
> ***Richard Dawkins. A Devil's Chaplin 2003 p.180[181]***

Sadly, orthodox medicine is not a house completely in order. It has repeatedly been shown to have large numbers of untested techniques in regular use. For instance, in a general medical ward in the United Kingdom in the mid 90's approximately 19% of interventions used were not evidence based[182].

Sleep Medicine has had a remarkable period of growth in EBM investigations since the mid 90s. Cohort studies such as the Wisconsin cohort[3] and the Sleep Heart Health Study[183] are now providing quality information about the prevalence, incidence, progression and associated consequences of SDB. The first systematic review and meta-analysis of CPAP treatment for OSAS found no reliable evidence of for the efficacy of CPAP[12]. That review was directly responsible for the halting of public funding for CPAP treatment in the UK. The role of the RCT in sleep medicine was rapidly expanded as a result, and has now reached the level where the construction of RCT-based meta-analyses are possible[111, 149]. Chapter 3 reports one such study.

Given the limited resources devoted to sleep medicine in socially-funded systems, it becomes imperative to maximise the benefit that can be achieved through the rational testing of the treatment methods at our disposal. The studies reported here help address the question of who should receive CPAP treatment, and the type of treatment they should receive.

CHAPTER 2
RANDOMISED CONTROLLED TRIAL OF CPAP IN PEOPLE WITH MILD TO MODERATE OBSTRUCTIVE SLEEP APNOEA

2.1 INTRODUCTION

Despite 6 published randomised placebo or conservatively controlled trials[90, 171, 184-187], the medium-term benefits of CPAP in the treatment of mild-moderate OSAS remain uncertain[149]. The trials are generally interpreted, by their authors, as showing some benefit for CPAP over placebo. However, the outcome measures that improve are not the same in each study, and the size of improvements tends to be small to moderate. In many of the studies, sleep apnoea symptoms scales show improvement, but these scales are not the same between research groups and always contain questions about witnessed apnoeas and snoring, which are factors that should be controlled by CPAP therapy. These scales are given little weight in the present literature review, because of this confounding. Findings from all 6 previous studies are chronologically reviewed below, and the sleepiness outcomes are meta-analysed in Chapter 3.

Previous Studies of CPAP Efficacy for Treating Mild-Moderate OSAS

The first RCT examining whether CPAP was an effective treatment for the daytime manifestations of mild OSA (AHI 5-15) was carried out with 16 patients (12 males) in Edinburgh[185]. All patients received both CPAP and an orally ingested placebo pill they were told "may aid upper airway function during sleep". The trial had an 11% dropout rate (2/18 enrolled). Mean objective compliance with CPAP was low, at 2.8 hours (SE 0.7 hrs) per night. Neither objective sleepiness (MSLT mean improvement = 0.1 mins, 95% CI = -1.0 to 1.3 mins), nor subjective sleepiness (ESS,

mean worsening =0.1 points, 95% CI = -2.1 to 1.9 points) were improved by CPAP compared to placebo, but individual symptoms scores, mental flexibility, and Hospital Anxiety and Depression (depression) scores were significantly better.

In the second published trial[187], 97 patients from an Ohio community sample of 'snorers' with mild to moderate sleep disordered breathing (AHI 5-30) and moderate daytime sleepiness, completed 8 weeks of treatment randomised to either CPAP or Conservative Therapy (CT). Both arms were given conservative management advice about sleep (sleep hygiene) and also weight loss advice from a dietician, if their BMI was greater than 29kg/m2. In addition, the CT group were given mechanical nasal dilators strips as a physical placebo (these strips do not lessen SDB). This trial has some unique strengths. It attempted to have equal numbers of males and females (52% male), and included a substantial number of African Americans (ca. 35%). Most trials have been almost exclusively included people of European descent and reflect the widespread clinical population bias toward males (70-90%). Objective compliance with CPAP was low, at 3.1 hours/night (44% of estimated sleep time). Only the vitality subscale of the SF-36 improved significantly more with CPAP than on CT. Responders to treatment were those who improved by at least an effect size of 0.50 in at least two of three domains: mood, energy/fatigue, and general health/functional status. Patients on CPAP were 2.72 times more likely to be responders (95% CI 1.18-6.58) than those on CT. The lack of additional improvement on CPAP might have been due to the marginal success of the Conservative Therapy to significantly reduce AHI by 3.7 over the testing period, although BMI was not reduced in either group.

In the third published trial[184], 34 Scots patients (21 males) with mild OSAS (AHI 5-14.9) were randomised to 4 weeks of either pill placebo or CPAP, and were then

crossed-over to the other treatment without washout. Post randomisation dropouts were low, at 8% (3/36). Objective CPAP use was low, at 2.8 (SD=2.1) hrs/night, and only 14/34 patients preferred CPAP to the placebo. After adjustment for placebo, CPAP improved symptoms, Epworth scores, 2 of 7 cognitive tasks, HADS depression, and 5 of 8 SF-36 subscales. However CPAP did not improve anxiety or objective wakefulness (Maintenance of Wakefulness Test). Of the 23 outcome measures used, 10 improved significantly more on CPAP than placebo. More highly compliant CPAP users (>2.5 hrs/night, n=17) improved more on 4 of the outcomes, and trended toward superior improvements in both the Epworth and MWT, than those who were less compliant. The size of the effects is probably not as large as the authors claim, as they have used a non-standard method for calculating effect sizes. Recalculated effect sizes using the baseline standard deviations presented in the paper yields effect sizes that are mostly small. Only for SF-36 Vitality and Health Transition were effect sizes moderate. The Epworth sleepiness scale, however, showed a large improvement of 3 points, when compared with a 3 point SD.

The largest mild-moderate (AHI 5-30) study was a 6-centre trial that followed 125 Spanish patients for 6 months of either CPAP and conservative therapy (CPAP+CT) or conservative therapy alone (CT)[186]. Patients were mostly male (86%), middle aged (54 yrs), and overweight (BMI 29±4). Objective CPAP compliance was the highest yet reported in this severity range, at 4.8 hrs per night. Given the trial duration, post-randomisation dropouts were low, at 12% (17/142). Only the OSAS-related symptoms scale improved more on CPAP+CT compared to CT. All other 20 outcomes, including the Epworth, MSLT, and blood pressure, showed no significant benefit for CPAP+CT compared to CT. On the full Functional Outcomes of Sleepiness Questionnaire, the difference at 6 months approached significance (p=0.06). This, the

largest trial of its type, offers no support for the widespread use of CPAP to treat patients with mild-moderate OSAS.

In a recent 2-centre Australian study, 28 of 42 patients (33% dropout), with mild-moderate OSAS (AHI 5-30), completed an 8 week crossover trial of CPAP and a pill placebo[171]. Patients were mostly male (24/28), middle aged (45± SD 11), and overweight (BMI 31±5 SD). CPAP use averaged 3.5 (SD 2.1) hrs/night. No outcome measure, except the symptoms questionnaire, improved more on CPAP than on placebo, including the Epworth Sleepiness Scale, MSLT, 24-hour monitored blood pressure, or quality of life. However, the symptom questionnaire used included questions addressing snoring, nocturnal choking, and witnessed apnoeas (this problem has been discussed above). This study does not offer convincing evidence that daytime function in mild-moderate OSAS patients is improved by CPAP.

The second Australian two-centre trial was a triple arm crossover, comparing CPAP, a Mandibular Advancement Splint, and pill placebo for 3 months each, with 2 weeks washout[90]. Eighty of 114 patients (79% Male, BMI 31±SE 0.6, aged 46.4±SE1.1) completed the entire 10.5 month trial (30% dropout). Compared to placebo, CPAP improved OSAS symptoms (see above for discussion), PVT lapses, the paced auditory serial addition task (PASAT 1.2 sec speed), the Profile of Moods States (POMS), the SF-36 mean score, and the Epworth Sleepiness Scale. Eleven other outcomes were not significantly better on CPAP than on placebo. The investigators also employed a relatively novel technique in this field, to deal with the multitude of outcome measures often used (something similar had also been done by Redline *et al.*[187]). Outcome measures were factor analysed and found to load on to the following dimensions; polysomnographic disease severity, symptoms (including sleepiness),

neurocognitive function, vigilance, and mood/quality of life. Improvements in most factors seemed to be greater on CPAP, than on placebo. However, only the improvement in polysomnographic indices was significantly greater. Effect sizes representing improvements from baseline within each arm, were summed. They indicated that CPAP was significantly better than placebo in reducing sleep disordered breathing, symptoms/sleepiness, and improving mood/quality of life. This study offers some evidence that mild-moderate OSAS can be treated successfully with CPAP. Results from this study need to be interpreted with some caution due to the dropout rate, which might have biased the results in favour of CPAP.

The study by Barnes and colleagues[90] is the most complex and comprehensive published in this field, and along with the two studies by Monasterio and colleagues[186] and Redline and colleagues[187], the largest. It currently offers the best extant evidence that CPAP has an effect in these patients beyond that offered by placebo, but the effect sizes are generally small.

All of these studies suffer from one important weakness that might have biased in favour of CPAP, in that none of them were double-blinded. The difficulty posed by double blinding in CPAP trials was discussed in Chapter 1. We had attempted to move closer to double-blinding by using sham CPAP as a control.

Considerations For the Design of the Present Study

One of the key factors potentially limiting CPAP's success in these studies has been suboptimal compliance, averaging 2.8[185] to 4.8 hours per night[186]. Studies have also had variable dropout rates, ranging from 8%[184] to 33%[171] that may have

biased the results in favour of CPAP. It might be possible to improve compliance in this type of trial by the addition of heated humidification to the CPAP circuit, as this has been shown to improve adverse upper airway symptoms and initial compliance[39, 178].

The use of different treatment comparisons [168, 170], including orally ingested placebo pills[90, 171, 184, 185] and conservative management advice[186, 187], might also have led to over or underestimation of the true benefits of CPAP. The best available choice of a placebo control is probably a sham CPAP device[12, 168], although no objective data exist to confirm this hypothesis. Sham CPAP has not been used with this subgroup before, and has never been combined with the advantageous crossover design, for any level of OSAS severity. Sham CPAP has advantages over an orally ingested pill in that it more closely controls for the negative effects of CPAP, such as sleep disruption and other CPAP interface problems. Conversely, it might also be expected to have a more potent positive placebo effect, due to the time and effort required by a participant and the researchers to establish a patient on a CPAP-like treatment. This investment in the therapy more closely resembles the interaction between patients and providers of CPAP therapy.

Another major benefit of the use of sham CPAP is that, unlike pills, the compliance level of an individual patient can be objectively monitored via the pressure switch in the device. However, the time-on-face compliance that is typically used to express compliance with CPAP treatment cannot be accurately measured in sham CPAP. Because of the use of internal flow resistors, used to manufacture a believable sham CPAP[169], the device cannot differentiate between the pressure of the mask when it is worn versus when it is not. Thus, the measured compliance with sham CPAP

reflects the time that a machine has been switched on, potentially overstating true compliance.

A crossover trial structure is arguably superior for investigating a chronic condition such as OSAS, especially in a general climate of important inter-individual differences in the effects of sleep disturbance on behaviour[188]. The combination of a crossover trial and a mechanical placebo, whose compliance levels can be objectively measured, allows the relative contributions of psychological and physical factors involved in the use of the device to be quantified and controlled for. A crossover trial should also incorporate a washout period between treatment arms. Two weeks is considered sufficient to washout the effects of CPAP[189].

The aim of this study was to assess the effectiveness of humidified nasal CPAP as a treatment for mild-moderate OSAS using a blinded, randomised, sham CPAP controlled crossover trial. Outcome measures were clinically relevant changes in daytime sleepiness, mood, quality of life, and psychomotor function. The effects of compliance could also be taken into account, as could inter-individual differences, via the use of mixed model analysis of variance.

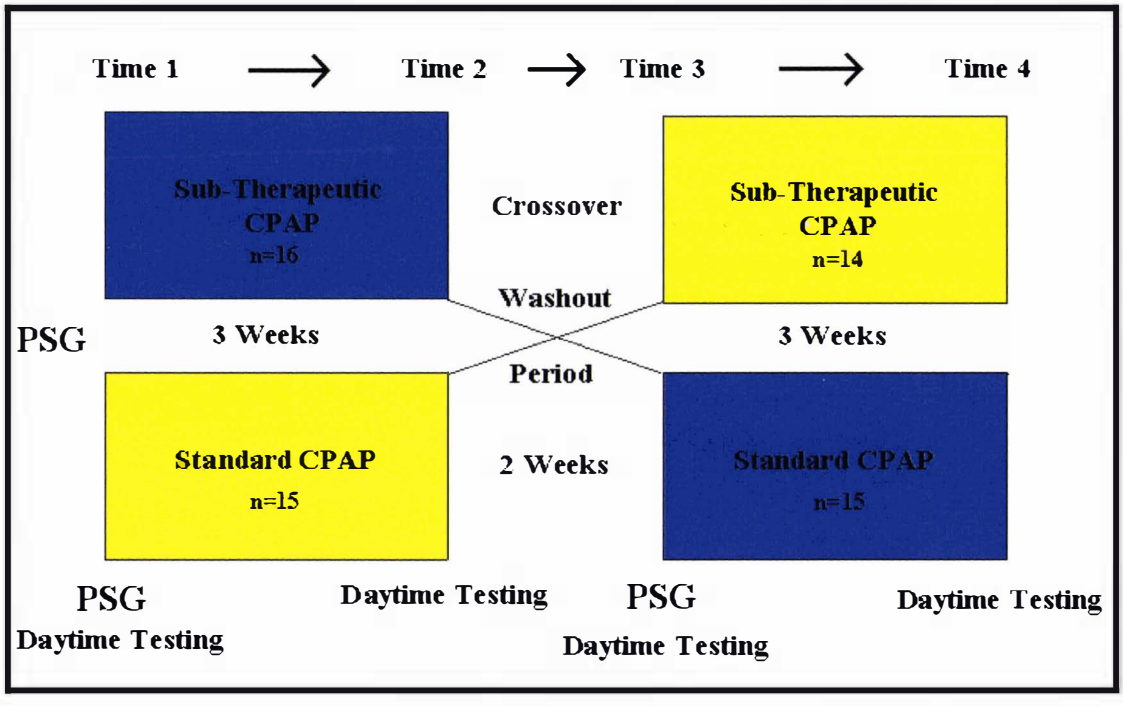
2.2 METHODS

Study Design

The protocol was approved by the Wellington Ethics Committee and devised by Dr Alister Neill. Each arm of the trial was three weeks long, with a two week washout period between arms to eliminate carry over effects. Allocation to treatment was achieved by drawing slips of paper without replacement from an urn. This was done

after the first day's testing had been completed, but before the evening polysomnography had begun. The urn contained 32 slips of paper indicating treatment sequence (either beginning on Sham CPAP, n=16 or standard CPAP, n=16). The numbers in these groups had been derived by simple coin tossing, which was done by Michinobu Imazu. Patients were informed that the study was "testing two different pressures of humidified CPAP". The investigator responsible for daytime study data collection (NSM) was blinded to treatment allocation. Deidre Sheppard, who supervised daytime data collection and collected some daytime data, was not blinded to treatment allocation after the first day. The sleep technicians who titrated CPAP or sham titrated placebo were also not blinded to treatment allocation and did interact with patients.

Figure 2.1: Pre Trial Flow Diagram



PSG=Polysomnogram. Note: Randomisation occurs before the first polysomnograms but after daytime testing at Time 1 Visit. Patient had already been diagnosed before entering the trial.

Power Calculation

It was calculated that 31 patients would be required to show a 2.5 point difference in improvement between placebo and CPAP in the Epworth Sleepiness Scale,

with 90% power, significance set at 0.05, and an inter-individual SD of 4.1 points in a cross-over design. The trial was thus powered to detect a moderate difference (2.5 / 4.1 = 0.61)[112, 184].

Pre-Trial Diagnostic Polysomnography

Diagnostic polysomnography (PSG) was undertaken either in a tertiary level sleep laboratory (WellSleep, Bowen Hospital), or in the patient's own home (S Series or P Series Sleep System, Compumedics, Melbourne, Australia), in accordance with accepted clinical and scientific methodology[190]. WellSleep is the clinical sleep investigation centre of the Department of Medicine, University of Otago at the Wellington School of Medicine and Health Sciences. WellSleep is listed by the Thoracic Society of Australia and New Zealand and the Australasian Sleep Association as an accredited sleep laboratory. Only private health system patients received at-home studies, as in-lab diagnostic polysomnography is the only diagnostic method funded by the public health system. At least 4hrs of sleep, including at least 30min REM sleep and 30 min sleep in the supine position, were required for a satisfactory diagnostic study. All studies were scored by Deidre Sheppard, Dr Angela Campbell or both.

Recruitment

Inclusion Criteria:

- 18 years of age or older
- English speaking
- Naïve to CPAP
- Diagnostic polysomnography demonstrated an apnoea hypopnoea index of 5-30/hr.
- Complaints of habitual snoring or nocturnal choking
  AND at least one daytime sleepiness symptom (daytime/evening napping, sleepiness whilst driving, never or rarely awakening refreshed)

OR Epworth Sleepiness Score ≥ 8.

- Written informed consent to be in the trial.

Exclusion Criteria:

- Extreme somnolence requiring immediate treatment;
- Shift-working (rotating or night work);
- Chronic sleep restriction (average Total Sleep Time ≤ 6 hrs/night);
- Current sedative, antidepressants, psychotropic or stimulant use;
- Alcohol intake > 3 standard units/24hr;
- Caffeine dependency;
- Upper airway surgery since the diagnostic sleep study;
- Clinically significant co-existing disease or additional sleep disorders (interpreted by Dr Neill).

Daytime Procedure

Testing began at 1230 hours on each of the four study days. All testing procedures were carried out at the same times of day on all study days (see Table 2.1 for times). From 1230 to 1330 patients were fitted with polysomnographic electrodes (electroencephalogram (C4-A1/C3-A2), left and right electro-oculogram, sub-mental electromyogram, and reference electrodes at Fpz and Fp1). All daytime data were collected by Deidre Sheppard and Nathaniel Marshall.

Table 2.1: Daytime Testing Timetable

Time of Day	Outcome Variable
1230	Patient Orientation.
	Placement of scalp and face electrodes
1330	Maintenance Of Wakefulness Test 1
1410	Epworth Sleepiness Scale
1430	Psychomotor Vigilance Task 1
1440	SF-36 General Health Survey
1530	Maintenance Of Wakefulness Test 2
1610	Hospital Anxiety and Depression Scale
1630	Psychomotor Vigilance Task 2
1640	Functional Outcomes of Sleepiness
1650	Electrodes Removed
1700	Patient Leaves (if visit 2 or 4).
	Blinded daytime data collector leaves (NSM)
Evening	Overnight CPAP or sham CPAP titration (if visit 1 or 3)

Within Trial Polysomnography and CPAP Titration

Overnight polysomnography was undertaken at the beginning of each arm for either CPAP titration, or to determine the effects of the placebo CPAP device. Comprehensive polysomnography included: electroencephalogram (electrodes placed at C4-A1 and C3-A2), left and right electro-oculogram, sub-mental electromyogram, abdominal and thoracic respiratory effort (piezoelectric bands), left/right leg movements, body position, pulse oximetry, oro-nasal airflow (by pressure transducer and thermistor), electrocardiogram, and where appropriate CPAP mask pressure. Studies were scored in 30 second epochs by either Deidre Sheppard or Angela Campbell, using accepted international criteria[191, 192].

The apnoea-hypopnoea index (AHI) was calculated by summing the total number of apnoeas and hypopnoeas and dividing by the number of hours of sleep. An obstructive apnoea was defined as a cessation of breathing for 10 seconds or more (<20% normal airflow), with ongoing respiratory effort. Without this effort, a 10 second apnoea or longer was defined as a central apnoea, while a combination of effort/no-effort defined a mixed apnoea. Hypopnoea was defined as greater than 50% reduction in the amplitude of at least 2 of the 3 respiratory channels (oro-nasal flow, thoracic, and abdominal effort, or nasal pressure) for at least 10 seconds, accompanied by either an arousal from sleep or ≥ 4% blood oxygen dip measured by fingertip pulse oximetry. Most previous RCTs have not used these more up-to-date criteria for defining hypopnoeas and have only monitored oro-nasal airflow and not nasal pressure. As thermistors are generally less sensitive than nasal pressure measurements, the estimates of AHI might be slightly higher when using nasal pressure measurements.

Continuous Positive Airway Pressure Systems

Both the standard CPAP and sham CPAP systems were based on the Fisher & Paykel HC221 device (Fisher & Paykel HC 221, Fisher and Paykel Healthcare, Auckland, New Zealand).

Therapeutic pressure was determined by manual titration by a sleep technician (Karyn O'Keeffe, Margo van den Berg, or Michi Imazu) and aimed to abolish apnoeas and hypopnoeas, and reduce respiratory related arousals. These effects were confirmed by a sleep physician (Dr Alister Neill). The sham CPAP[169] device, developed and tested in Wellington for this trial by University of Otago medical student Nick Burgess, was set to 8cm H2O, but actually delivered <1.0cm H_2O, due to placement of a resistor

at the pump outlet, and a modified mask with extra holes drilled around the exhalation port (Mirage, ResMed, Sydney, Australia). The mechanical CPAP placebo created conditions that were identical to CPAP at 8cmH2O in terms of noise, mask temperature, mask humidity, and air flow through the exhalation port.

Outcome Measures

Sleepiness

The Epworth Sleepiness Scale (ESS)[193] is in widespread use in sleep medicine (see Appendix 1). Factor analysis has indicated that it is a unitary measure of subjective sleepiness. It asks 8 questions about the likelihood of dozing in everyday situations. Each question can be answered 'None' (0), 'Slight' (1), 'Moderate' (2) or 'High' chance of dozing (3). The resulting scale is scored between 0 and 24, with scores greater than 10 generally regarded as high and above 15 as very high (see Appendix I). New Zealand population norms for this questionnaire have recently become available and in the Non-Māori population the average Epworth score is about 6.5[11].

Wakefulness

A modified Maintenance of Wakefulness Test was used to measure objective daytime wakefulness (modMWT)[194-196]. The MWT is an electrophysiological test of wakefulness and is regularly employed in clinical trials of treatments for OSAS (see Appendix 1). Patients are left alone propped up in bed at about 70° to the horizontal, with their street clothes on, and are instructed to remain awake as long as possible. Then the lights are turned out. The test is normally repeated 4 or 5 times throughout the day (typically 10am, noon, 2pm, 4pm, and sometimes 6pm). Sleep/wake is scored via live interpretation of the primary EEG channel (C4-A1). Three consecutive 30 epochs of stage 1 sleep, or any epoch of another stage of sleep (2, 3, 4 or REM), are taken to

indicate unequivocal sleep onset. If the patient is not asleep after 40 mins, the test is terminated and the sleep latency is recorded as being 40 mins. The sleep latencies are averaged to give the mean sleep latency out of 40 mins. Normative data from the United States and Australia indicate that normal mean latencies are around 35-38 minutes[195, 196].

To enhance patient recruitment and study completion, we abbreviated the Maintenance of Wakefulness Test, as described in an abstract by Banks and colleagues[197], to include two 40-minute tests which were started at 1:30pm and 3:30pm, rather than the typical 4 naps. The test used here is therefore termed the modified Maintenance of Wakefulness Test (ModMWT), to avoid confusion.

Quality of Life

The Functional Outcomes of Sleepiness Questionnaire (FOSQ)[198] consists of five subscales scored out of four points, that sum to an overall scale out of 20. The five subscales ('General Productivity'- 8 items, 'Social Outcome'- 2 items, 'Activity Level'- 9 items, 'Vigilance'- 7 items, and 'Intimacy and Sexual Relationships'-4 items) aim to measure quality of life, as it is affected by disorders of excessive sleepiness such as OSAS. The FOSQ scale was employed in this study without the sexual functioning subscale, in order to avoid differences in comparison between patients who were and were not in sexual relationships. Some normative data are available from Weaver's original introduction of the scale[198], but a larger sample of non-sleepy individuals is required to determine variability between genders, ages and ethnic groups.

The SF-36 is a reliable, widely validated 36 item abbreviated version (SF-short form) of the Medical Outcomes Survey (MOS), containing 8 independent sub

scales[199-201]. As a general measure of health-related quality of life, the SF-36 aims to measure multiple dimensions of health that are generally applicable across disease groups. However, this approach requires a trade-off between the breadth and depth of enquiry, and the expected time of completion of 5-10 minutes. Each of the SF-36's subscales are scored between 0 and 100. These independent subscales are Physical functioning (PF-10 items), Role Physical (RP-4 items), Bodily Pain (BP-2 items), General Health Perceptions (GH- 5 items), Vitality (VT-4 items), Social Functioning (SF- 2 items), Role Emotional (RE- 3 items), Mental Health (MH- 5 items) and Reported Health Change over a 1 year period (1 item)[200-203]. The final measure was not employed due to the short time frame of this trial. These subscales primarily load onto and can be combined into Physical and Mental Component summary scales[202], however we do not report those scales here. New Zealand normative data is available[204].

Mood

The Hospital Anxiety and Depression Scale (HADS) [205] aims to screen for pathological anxiety and depression in clinical populations. The scale is bidimensional, although the dimensions are correlated[206]. The HADS consists of 14 questions, 7 related to anxiety, and 7 to depression. Each of the 14 questions has 4 options that are scored between 0 and 3. Thus each of the scales 'Anxiety' and 'Depression' range between 0 and 21. New Zealand normative values do not seem to be available, however in a relatively small normative general population sample (n=199), aged 18-65, Spinhoven *et al.*[207] report an average Anxiety score of 5.1 (SD 3.6) and average Depression score of 3.4 (SD 3.3).

Psychomotor Performance

The Psychomotor Vigilance Task (PVT)[154] is a handheld device that tests simple primary reaction times. Two ten minute tests were administered per day, with a 1 minute practice session before each test. The device briefly feeds back performance to patients by giving reaction times in milliseconds. Mean reaction times, lapses (reactions longer than 500ms), and reaction errors (particularly false starts) were measured. Generally, reaction times of around 260ms might be expected in a group of healthy middle-aged people, but accurate reference values are not available. The PVT is sensitive to the effects of sleep restriction and fragmentation and this seems to follow a stable pattern within individuals[154, 188, 208]. PVT Lapses (responses slower than 500ms) have been shown to be sensitive to the effects of CPAP in a randomised controlled trial, although the observed improvement on CPAP was partially due to a slight worsening on placebo[90].

Preference and Compliance

Compliance was objectively measured by an internal pressure sensor and expressed as the average number of hours use per night (HC 221, Compliance Maximiser release 1.01 Fisher & Paykel Healthcare, Auckland New Zealand). The amount of time that the mask was worn on the face at therapeutic pressure was measured during the CPAP arm. Compliance whilst on sham CPAP could only be ascertained by machine running time.

At the end of the trial, patients were asked via questionnaire which treatment was preferred, and which they thought was better for their sleep.

Data Management and Analyses

Data were independently double entered by two investigators (Deidre Sheppard and Nathaniel Marshall). Significance and size of treatment effects were evaluated in SAS (SAS institute v.8) and SPSS (v.12). All data analyses and interpretation were done by Nathaniel Marshall with advice from a biostatistician, Gordon Purdie (Wellington School of Medicine). Data analyses were not subject to treatment blinding. Results for the modified MWT and PVT were averaged across each day. Continuous data were tested for normality using the Kolmogorov-Smirnov test and visual checking of histograms under normal curves. Paired sample t-tests were used to detect changes from baseline sleep variables during sham CPAP treatment. Chi-squared tests were used to test for patient preference between CPAP and placebo. The Wilcoxon test was used to test between the CPAP and sham CPAP conditions in non-continuous data.

Mixed model analyses of variance were employed to investigate treatment effect differences in continuous, normally distributed outcome improvements. CPAP compliance was split into a dichotomous variable at 4 hours per night for the mixed models. Treatment, order of treatment (1st or 2nd arm), compliance, and interactions of treatment by compliance and treatment by order, were entered into the mixed models as fixed effects, and participants as a random effect. Main effects, and any interactions, were regarded as statistically significant when $p<0.05$. Main effects were interpreted as favouring treatment when CPAP was significantly better than sham CPAP AND the pattern, from the treatment by compliance interaction, indicated a clear benefit of high compliance on CPAP over both the sham CPAP combinations.

Whenever a mean reduction/improvement in an outcome measure is quoted, it is an estimated mean produced by a mixed model. These closely resemble the arithmetic mean, but do differ slightly because of the model estimation procedure.

Effect sizes were used to quantify the magnitude of the treatment effects and were calculated by dividing the mean effect of CPAP (placebo adjusted) by the standard deviation of the outcome measure at baseline. Small effect sizes are between 0.20 and 0.50, medium 0.50 and 0.80, and large effect sizes >0.80[112].

2.3 RESULTS

Participants

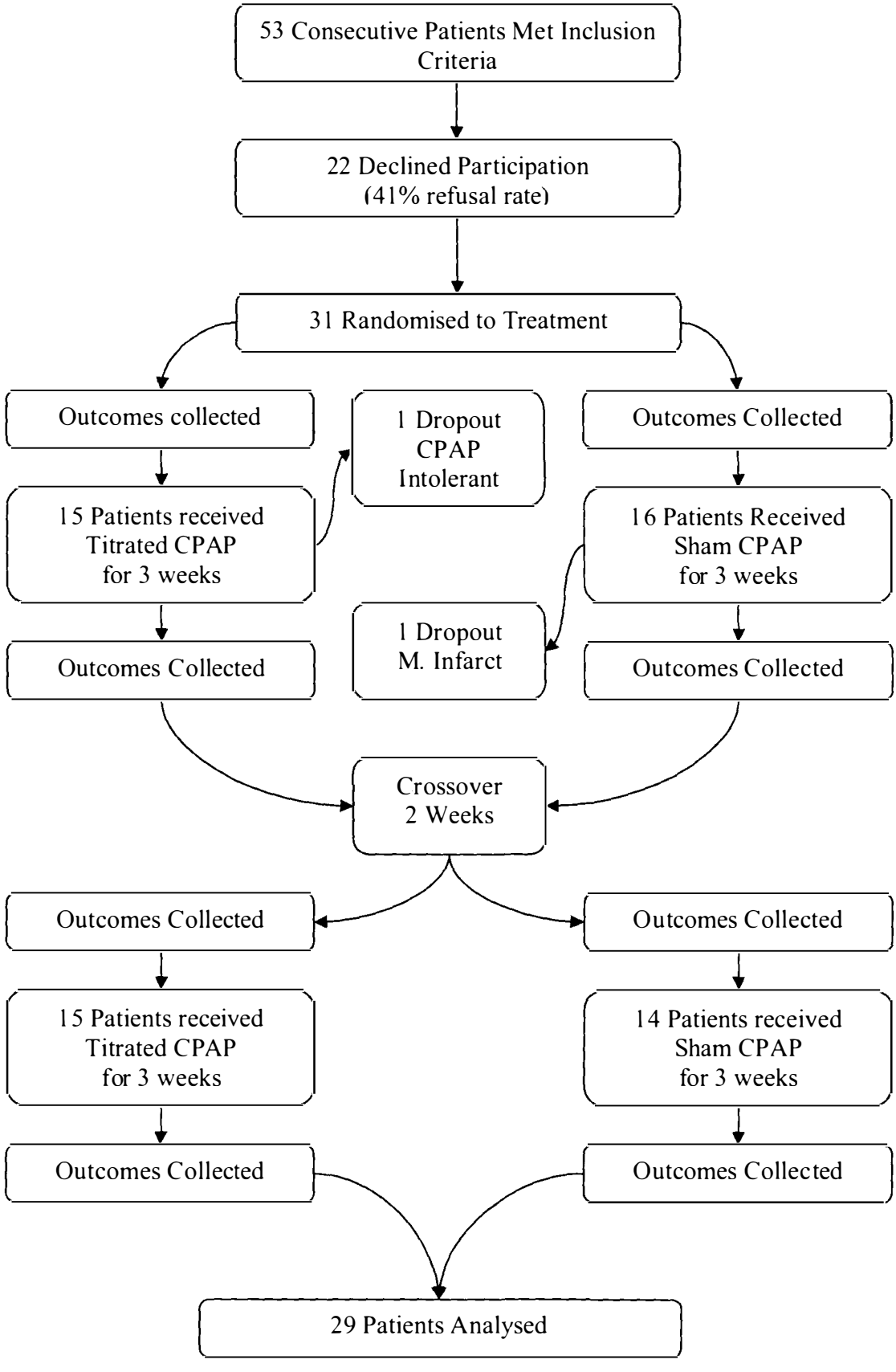
Thirty-one of the 53 (58.5%) eligible patients agreed to participate. Sixteen patients said they did not have enough time or were uninterested in being enrolled in the study, 3 had moved outside or lived outside the enrolment zone, and 3 decided to use treatments other than CPAP. Non-participants were not significantly different in terms of age, gender, body mass, referral source, Epworth Score, sleep efficiency, or arousal indices. Non-participants had lower average apnoea hypopnoea indices (see Table 2.2). But this difference was probably not clinically relevant. These patients are also comparable to a general clinical sample from WellSleep reported by Mihaere[33] where Māori made up approximately 14% of the sample, women 24%, and the average age was about 48 years.

Table 2.2: Characteristics of Participants and Non-Participants

Patient Characteristic	Participants	Non-participants	p for difference
Gender (Female)	7/31 (23% female)	7/22 (32% female)	0.45†
Referral Source	12* Public 19 Private	10 Public 12 Private	0.62†
Body Mass Index kg/m^2	31.2 (SD 5.6)	31.2 (SD 6.8)	0.97
Age (Years)	51 (SD 11)	46 (SD 11)	0.10
Epworth Score (/24)	12.0 (SD 4.4)	9.6 (SD 5.1)	.08
Sleep Efficiency (%)	84 (SD 9)	80 (SD 13)	.17
Arousal Index (/hr)	26.9 (SD 8.3)	23.6 (SD 7.4)	.14
AHI	22.4 (SD 8.0)	15.9 (SD 6.0)	**<0.01**

* also includes one patient referred from a previous research project where OSAS was diagnosed. † Chi Square test SD=Standard Deviation. AHI- Apnoea Hypopnoea Index.

Figure 2.2: Post Trial Flow Diagram



Randomisation and Retention

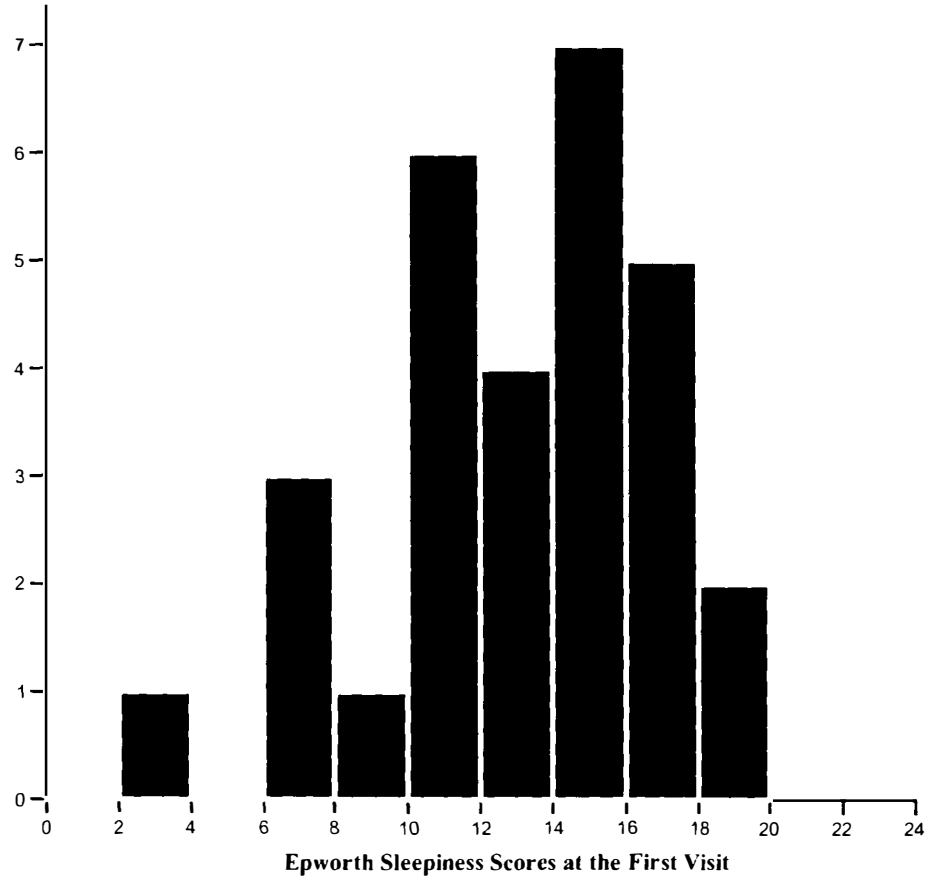
Thirty one patients were randomised to treatment, with 15 beginning on CPAP and 16 on sham CPAP. Two patients dropped out post randomisation (6.5%), both in the first arm (see Figure 2.2). The first was a middle-aged (56 years) male, who was very sleepy (ESS=16), overweight (BMI=27.7), and had moderate SDB (AHI= 28.7). This patient suffered a non-fatal myocardial infarction during the placebo treatment arm. It was decided by the medical director (Dr Alister Neill) to remove this patient due to possible effects on the outcome measures being examined. The other patient was an overweight (BMI=28.5), moderately sleepy (ESS=9), retired male (64 years), with moderate SDB (AHI=27/hr). The patient was intolerant of the noise of the CPAP device, and was not prepared to continue the trial (self withdrawal). Analyses are presented on the remaining 29 patients (see Table 2.3), 14 beginning on CPAP and 15 on placebo, as the two dropout patients presented for only one of the 4 data-collection points and could offer no meaningful data. The analyses were being conducted under the "intention to treat" principle, as some of the remaining patients did not effectively use CPAP, but were still included in all analyses.

Table 2.3: Characteristics of Patients Completing the Trial

Patient Characteristic	Mean or Median	Measure of Variability
Gender (Female)	7/29 (24% female)	
Referral Source	11 Public* 18 Private	
Body Mass Index kg/m^2	31.5	SD 6.0
Age (Years)	Median 50.6	range 25-67
Epworth Score (/24)	12.4	SD 4.1
Ethnicity	22 Exclusively European/Pakeha 7 Māori and/or Polynesian	
Diagnostic Study Type	11 In Lab. 18 At home	
Sleep Efficiency (%)	84.5	SD 8.9
Arousal Index (/hr)	25.1	SD 7.1
Apnoea Hypopnoea Index	21.6	SD 7.5
Prescribed CPAP Pressure	7.3 cmH_2O	Range 5-10cmH_2O

* also includes one patient referred from a previous research project.

Figure 2.3: Baseline Sleepiness Distribution



Effects of Sham CPAP and CPAP on Sleep

Table 2.4 indicates that the treatment of SDB by CPAP was successful. It also indicates that sham CPAP did not alleviate SDB, and that it therefore is a valid placebo treatment.

Sham CPAP is not a benign device, (see Table 2.4). It has a number of significant deleterious effects on sleep quality, including a shift away from slow wave sleep toward stage two sleep, and a doubling of sleep latency. These effects might have helped cause a trend (non-significant) toward worsening in maintenance of wakefulness latencies on placebo. However, sham CPAP is a successful placebo control for CPAP because it mimics the sleep disruption caused by CPAP whilst not reducing the indices of sleep disordered breathing, including the AHI, SaO_2 desaturations and the arousal index. The figures in Table 2.4 may need to be interpreted with some caution as they represent the first nights that all patients used sham CPAP and in about half of cases the first use of any CPAP device. Eighteen of the patients had also undergone their diagnostic polysomnography recording at home and thus were spending their first night in the sleep laboratory environment. Thus, some of the effects on sleep quality of CPAP, and sham CPAP might be due to a first night effect. Subsequent nights of sham treatment might differ slightly or substantially from that presented below, as patients habituated to the device and also returned to their own sleeping environments.

By contrast, CPAP significantly improved the AHI, arousal index, lowest oxygen desaturation, and mean desaturation (see Table 2.4). In common with sham CPAP, it worsened sleep efficiency and sleep latency, and shifted sleep architecture

into stage 2 and away from stages 1 and 4. Importantly, it successfully reduced the harmful indices of SDB, as it is designed to. The data presented include time where CPAP was not at full pressure, and thus SDB might not yet have been abolished. Presenting data from much smaller time periods where therapeutic pressure had been achieved may give a less accurate description of effects, due to confounding from changes in sleep architecture during the latter parts of the night, where more REM and less slow wave sleep are normally seen.

The titration night process might also explain the worsening in sleep latency and sleep efficiency on both CPAP and sham nights. Measuring sleep a week or two into each treatment might have given a more accurate picture of the effects of both devices on sleep, but unfortunately this was not achievable.

Table 2.4: Effects on Sleep of Sham CPAP and CPAP

Sleep Variable	Diagnostic Night	Sham CPAP Night	CPAP Night	Between Treatments (p)
Sleep Efficiency (%)	84.5 (8.9)	79.2 (11.7)*	75.2 (11.6)**	0.07
Sleep Length (Mins)	364 (74)	357 (69)	340 (66)	0.12
Sleep Latency (mins)	10.1 (11.3)	21.3 (16.6)**	21.6 (19.6)*	0.95
Arousal Index (/hour)	25.1 (7.1)	22.2 (9.8)	13.8 (6.6)**	<0.01
SaO2 lowest point (%)	85.7 (3.7)	85.1 (6.5)	87.8 (5.2)*	<0.01
Mean SaO2 desat (%)	3.6 (1.1)	3.7 (1.7)	2.57 (1.07)**	<0.01
Total AHI (/hour)	21.6 (7.5)	22.0 (14.5)	7.9 (6.7)**	<0.01
Supine Sleep (mins)	127 (88)	175 (99)*	186.5 (108.9)	0.44
Stage 1 Sleep (%)	10.8 (5.3)	9.0 (4.7)	7.1 (3.2)**	0.03
Stage 2 Sleep (%)	46.3 (8.3)	56.5 (7.7)**	54.9 (9.7)**	0.27
Stage 3 Sleep (%)	13.1 (6.4)	12.1 (5.1)	13.0 (5.7)	0.41
Stage 4 Sleep (%)	10.2 (6.4)	5.4 (4.0)**	6.2 (6.4)**	0.38
Stage REM (%)	19.6 (6.3)	17.0 (5.6)	18.8 (5.2)	0.06

Legend: Values indicate the mean (and standard deviation) *p<0.05; **p<0.01 indicate significant differences within patients compared to the diagnostic night. All variables were approximately normally distributed.

CPAP Pressure and Compliance with CPAP Therapy

Titrated CPAP pressure ranged between 5 and 10cmH2O. Average compliance with humidified CPAP was objectively measured at 4.9 hours per night, and ranged between 0 and 8.4 hours per night (see Figure 2.4). Eighteen patients (62%) used the device more than 4 hours per night. Sixteen patients (55%) used the device on average more than 6 hours per night. Average run time on sham CPAP was also 4.9 hours per night (range 0-8.32 hrs), and was highly comparable to CPAP (see Figure 2.5). Compliance with CPAP treatment was strongly positively correlated with compliance on the Placebo treatment (Pearson's $r=0.861$ $r2= 0.741$, $p<0.001$ see Figure 2.6). Compliance during the CPAP arm was significantly negatively correlated with CPAP pressure (Pearson's $r=0.527$, $r2=0.278$, $p<0.01$ see Figure 2.7). Compliance with placebo was also negatively correlated with the unrelated pressure that CPAP was set to in the opposite arm (Pearson's $r=0.422$, $r2=0.178$, $p<0.05$ see Figure 2.8). This relationship is similar to that for compliance and pressure within the CPAP arm. For the purposes of mixed modelling, high compliance was set at ≥ 4 hours per night. By this criterion one patient was a high complier on CPAP and then a low complier on placebo and one followed an inverse pattern (see Figure 2.6). Both were assigned compliance levels in the models based solely on their CPAP compliance.

A potential anomaly in the compliance comparison is that there were two patients who were moderately compliant with placebo (ca.3.5hrs/night) but almost totally non-compliant with CPAP. Their apparent higher compliance with placebo might be an artefact of the inability of the HC221 device to identify whether the mask

was being worn during the placebo condition. Indeed, the potential exists that the compliance with placebo was systematically overestimated.

Figure 2.4: Compliance with CPAP Therapy

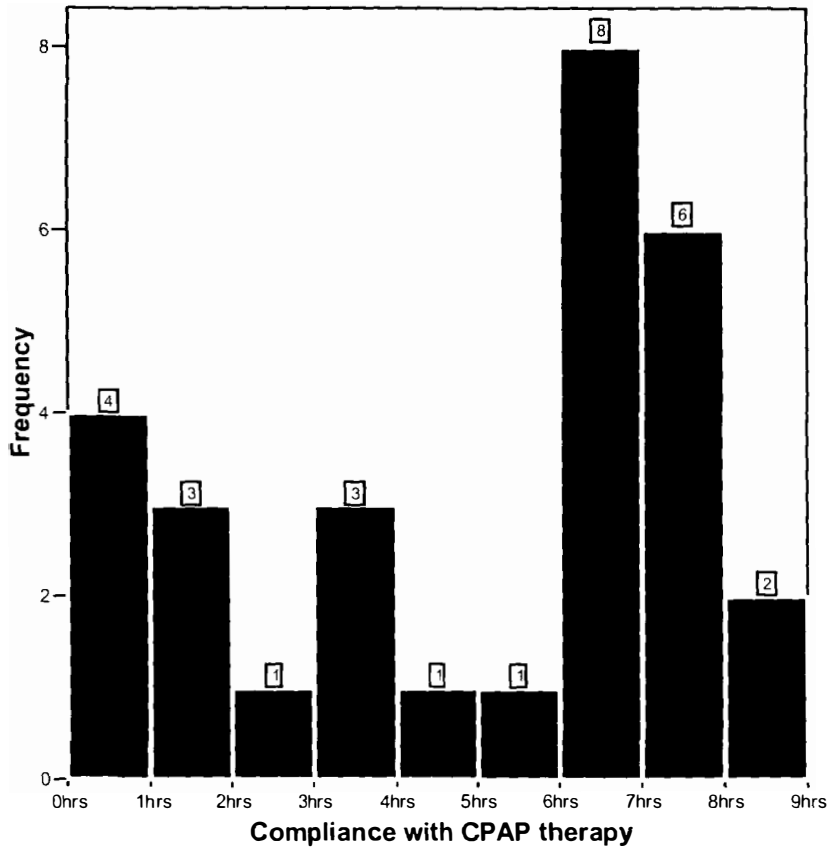


Figure 2.5: Compliance with Sham CPAP Therapy

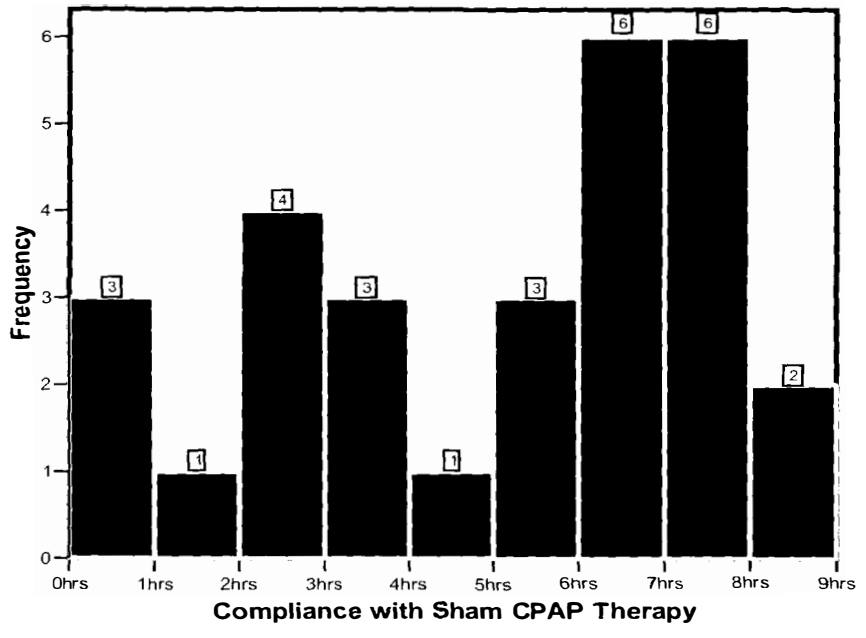


Figure 2.6: Scattergram Showing Correlation Between Sham and Standard CPAP Compliance

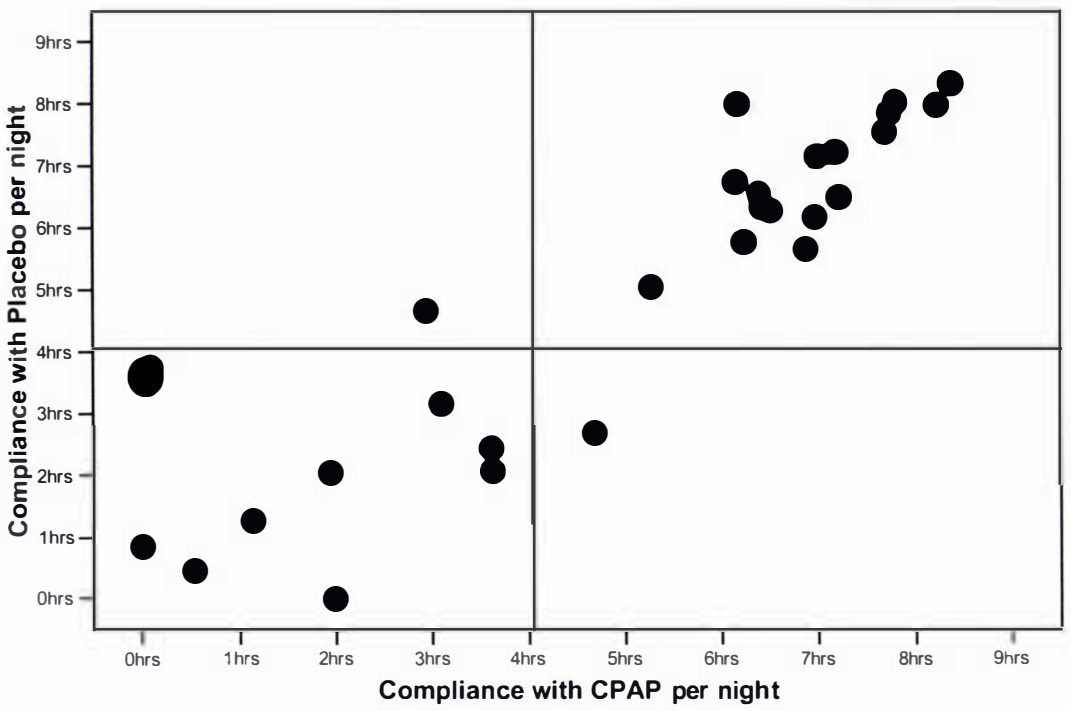


Figure 2.7: Scatter-gram Showing Correlation Between Compliance and Pressure on the CPAP arm.

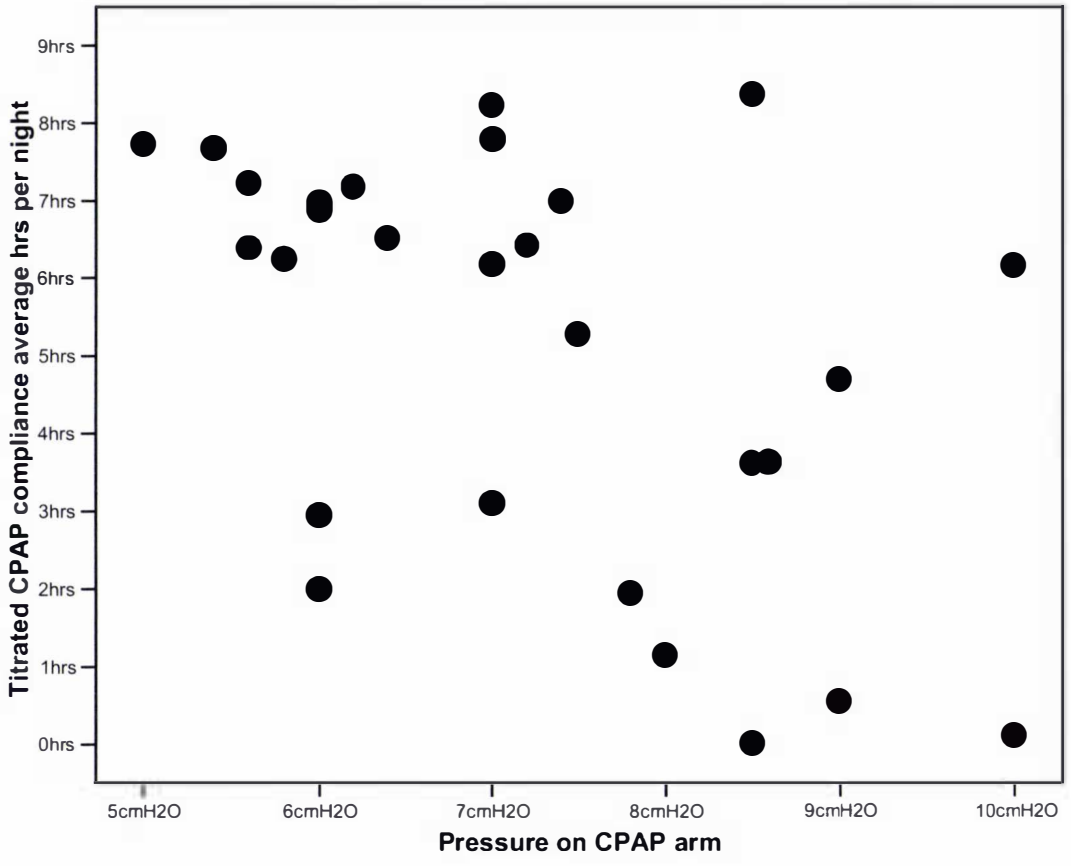
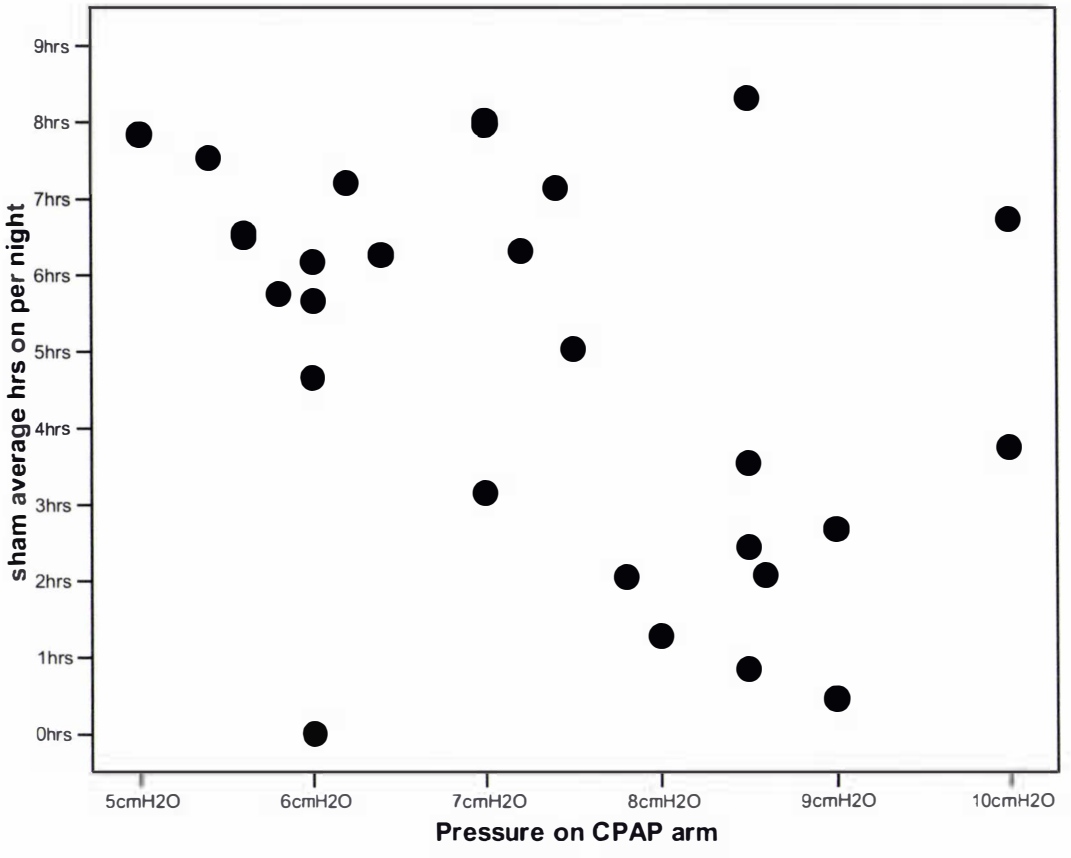


Figure 2.8: Scatter-gram Showing Correlation Between Sham CPAP Compliance and Pressure on the Unrelated CPAP Arm.



Main effects of Treatment

Six of the twenty outcome measures failed the Kolgomorov-Smirnov test ($p<0.01$ compared to normal). However 5 of the 6 had tolerably normal distributions and had failed the K-S test due to one or two outliers. Only the Role Emotional Subscale of the SF-36 was non-normal, largely because it was also non-continuous. It was analysed using a Wilcoxon rank sum test. All of the other outcome measures were deemed suitable for analysis with mixed model ANCOVA.

The Mixed Model ANCOVAs (see Table 2.6) were constructed identically with the following fixed effects and interactions: treatment (CPAP vs. sham CPAP), order (1st vs. 2nd arm of trial), compliance (High vs. Low Compliers), treatment by order (differential effects of treatment depending on arm in which received), treatment by compliance (dose-response relationships should be evident if CPAP is effective).

The main effects of CPAP and placebo on all outcome measures are presented in Table 2.5. No carry-over effects were observed. Humidified CPAP improved subjective sleepiness (measure by Epworth Sleepiness Scale). The effect size (0.58) was moderate (estimated mean improvement = 2.39 points, 95% CI=0.6 to 4.2, baseline SD 4.1). There was evidence of a simple dose-response relationship, with high-compilers (≥ 4 hours /night) improving by an estimated mean 2.6 points (95% CI=1.5 to 3.7) compared to the low-compliers 0.7 points (95% CI= -0.8 to 2.1, p for difference=0.04). The effects of treatment and compliance were additive in the expected direction (see Figure 2.9), high compliance on CPAP (estimated mean improvement =3.9 points, 95% CI 2.3 to 5.5) was better than both high compliance on placebo (estimated mean improvement=1.2 points, 95% CI=-0.3 to 2.9, p for difference=0.02) and low compliance on placebo (estimated mean worsening=0.4 points, 95% CI= -2.4 to 1.7, test

for diff $p<0.01$). High CPAP use also tended to be better than low use of CPAP, but this was not significant (estimated improvement=1.7 points, 95% CI= -0.3 to 3.8, p for difference=0.09).

Table 2.5: Summary of Effects of CPAP and Sham CPAP on Outcomes

Outcome Measurement (and Maximum Score or Standard Unit size)	Baseline Mean (SE)	Placebo Mean (SE)	CPAP Mean (SE)	Net Effect (95% CI)	Effect Size	CPAP Usage (p)
Epworth Scale (24)	12.5 (.8)	12.0 (.7)	9.7 (.7)***	**2.4 (0.6-4.2)**	**0.58**	**0.04**
ModMWT (latency, mins)	20.9 (2.5)	17.9 (2.0)	23.1 (2.0)	5.2 (-0.6-11)	0.40	0.09
FOSQ Total (16)	12.6 (.3)	13.3 (.3)	13.6 (.3)**	0.3 (-0.5-1.1)	0.14	0.96
-Activity (4)	3.0 (.1)	3.3 (.1)*	3.3 (.1)**	0.1 (-0.1-0.3)	0.02	0.82
-Social Outcomes (4)	3.2 (.1)	3.5 (.1)*	3.7 (.1)**	0.2 (-0.2-0.5)	0.17	0.20
-Vigilance (4)	3.0 (.1)	3.1 (.1)	3.2 (.1)	0.1 (-0.2-0.3)	0.13	**0.05**
-General Product. (4)	3.3 (.1)	3.4 (.1)	3.4 (.1)	0.0 (-0.2-0.2)	0.13	0.65
SF-36 (all /100)						
- Mental Health	75 (3)	80 (2)*	77 (2)	-3 (-10-3)	-0.22	0.74
- Bodily Pain	73 (4)	75 (4)	68 (4)	-7 (-17-3)	-0.31	**0.04**
- Social Functioning	79 (4)	85 (5)	78 (5)	-8 (-21-6)	-0.36	0.26
- Vitality	44 (3)	53 (3)**	53 (3)**	-1 (-10-9)	-0.04	0.12
- Role Emotional	80	89	81	NS	N/A	N/A
(Inter Quartile Range)	(67 to 100)	(89 to 89)	(64 to 81)			
- Role Physical	63 (8)	82 (7)**	66 (7)	-16 (-35-3)	-0.38	0.85
- Physical Functioning	82 (3)	80 (2)	81 (2)	1 (-3-6)	0.09	**0.02**
- General health	74 (3)	76 (2)	76 (2)	0 (-6-7)	0.03	0.46
HADS- Anxiety (21)	6.8 (.7)	5.7 (.4)**	6.5 (.4)	-0.8 (-1.9-0.4)	-0.20	0.52
HADS- Depression (21)	4.2 (.5)	3.5 (.3)*	3.8 (.3)	-0.3(-1.2-0.6)	-0.10	0.29
Psychomotor Vigilance Task –						
- Mean RT (ms)	264 (5)	259 (5)	266 (5)	-7 (-20-7)	-0.15	0.16
- Lapses (>500ms)	1.3 (.3)	1.0 (.4)	1.3 (.4)	-0.4 (-1.4-0.7)	-0.18	0.15
- Errors	2.8 (.5)	3.3 (.7)	3.2 (.7)	0.1 (-1.9-2.0)	0.02	0.27

Legend: Measures (maximum score or standard units) with means (SEM) at baseline and for the effects of CPAP and placebo compared to baseline. * $p<0.05$ ** $p<0.01$ *** $p<0.001$ compared to baseline.. Negative effect sizes indicate Placebo outperformed CPAP. CPAP usage column indicates simple dose response relationships. Only the Vitality subscale of the SF-36 and the PVT Mean RT exhibited any order effect. Interactions are discussed in the text. Means and SEMs estimated by the mixed models. Baseline values are from the first day of testing. SF36 Role Emotional data were non-normal and non-continuous and were compared using the Wilcoxon test. The interquartile range is listed where standard error is listed for other variables. NS = Not significant (Wilcoxon).

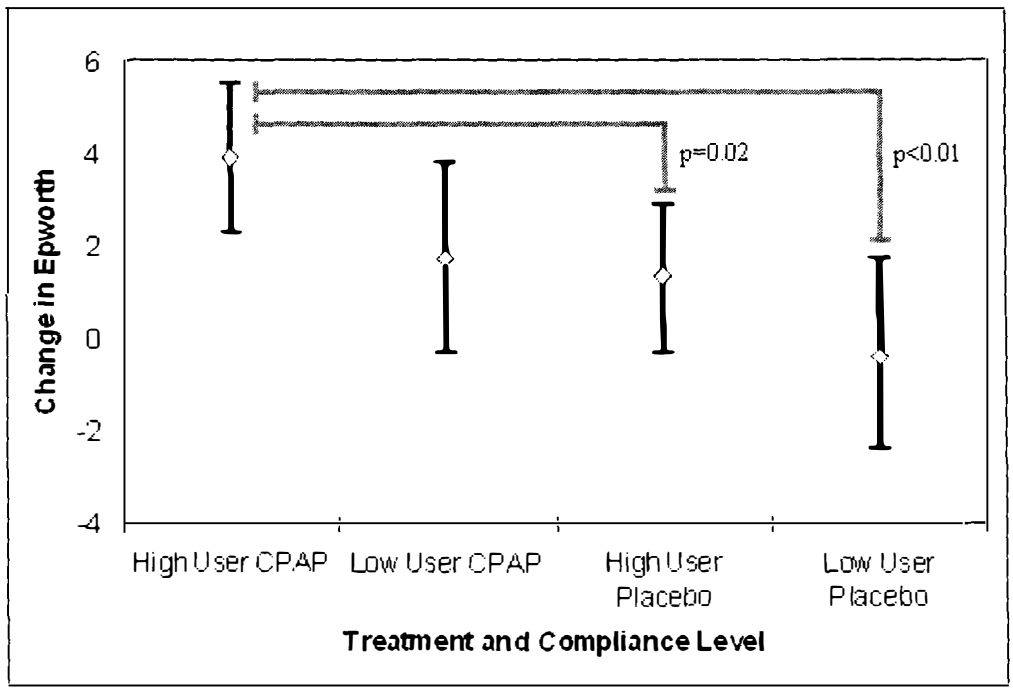
Abbreviations: ModMWT= Modified Maintenance of Wakefulness test. FOSQ= Functional Outcomes of Sleepiness Scale, SF-36 Short form 36 Question version of the Medical Outcomes Survey, HADS= Hospital Anxiety and Depression Scale, RT= Reaction time.

Table 2.6: Mixed Model ANCOVAs for Outcome Measures

Dependant Variable	Fixed Effects in Model	Significant Main Effects
Epworth	Treatment, Order of Treatment, Compliance (high/low) Treatment by Order Treatment by Compliance	Treatment (95% CI= 0.5 to 4.2, p=0.012) Compliance (95% CI= 0.1 to 3.8, p=0.040)
Modified MWT	As above	None
FOSQ Total	As above	None
-Activity	As above	None
-Social Outcomes	As above	None
-Vigilance	As above	None
-General Product.	As above	None
SF36 - Mental Health	As above	None
SF36 Bodily Pain	As above	Compliance (95% CI= 0 to 20, p=0.040)
SF36 Social Functioning	As above	None
SF36 Vitality	As above	Order (95% CI= 2 to 21, p=0.017)
SF36 Role Emotional	N/A	N/A
SF36 Role Physical	As above	None
SF36 Physical Functioning	As above	Compliance (95% CI= 1 to 11, p=0.023)
SF36 General health	As above	None
HADS- Anxiety	As above	None
HADS- Depression	As above	None
PVT Mean RT	As above	Order (95% CI= 4 to 30, p=0.014) Treatment*Compliance (F=4.92, p=0.031)
PVT Lapses (>500ms)	As above	None
PVT Errors	As above	None

Main Fixed Effects are only listed where $p<0.05$. Sf-36 Role Emotional subscale was analysed with a Wilcoxon rank sum test.

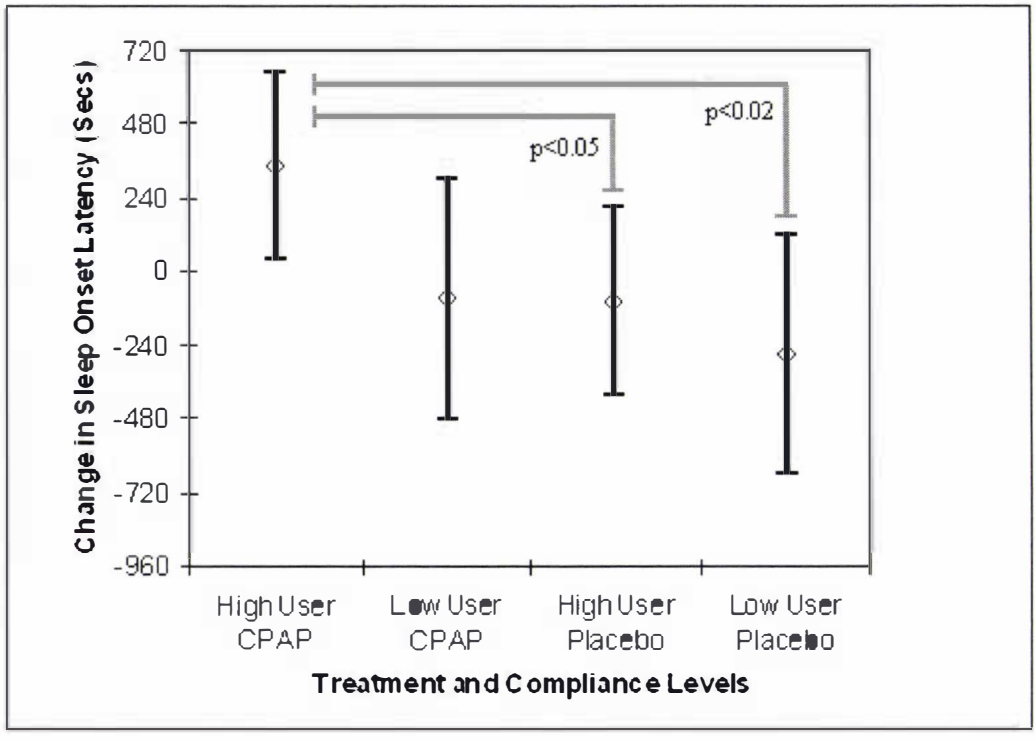
Figure 2.9: Interactions Between Treatment and Compliance in the Epworth Sleepiness Scale



Objective wakefulness (modMWT) exhibited a net trend toward improvement on CPAP (estimated mean improvement=5.2 mins, 95% CI= -0.6-11) with a small effect size of 0.40 (5.2 mins (95% CI= -0.6 – 11 mins). Because post hoc power calculations cannot yet be calculated for mixed models[209], power calculations were based on t-tests with 80% power and a significance level of 0.05 with standard deviations calculated from the first day of the trial. Once differences are expressed as effect sizes, it is possible to calculate that an effect size needs to be around 0.54 (i.e. moderate) in order to have an 80% chance of being found significant at the 0.05 level. Reversing the same methodology gives a figure of 41 patients to show an effect size of 0.40. There was some evidence of the expected dose-response relationship in the MWT (see Figure 2.10). In highly compliant users, wakefulness improved by an estimated mean 5.7 mins on CPAP (95% CI= 0.7 to 10.8 mins), which was significantly better than when those same high compliers received sham CPAP (estimated worsening =1.6 mins, 95% CI= -6.7 to 3.5 mins, p for difference=0.046), and better than poor compliers

on sham CPAP (estimated worsening =4.5 mins, 95% CI= -11 to 2 mins, p for difference=0.016). Highly compliant CPAP users tended to improve more than less compliant CPAP users (estimated worsening =1.5 mins, 95% CI -8 to 5 mins), but this difference was not significant (p for difference=0.086).

Figure 2.10: Interactions between Treatment and Compliance in the Modified Maintenance of Wakefulness Test



Mood (Hospital Anxiety and Depression Scale), quality of life (SF 36, The Functional Outcomes of Sleep Questionnaire) and simple reaction times (Psychomotor Vigilance Task) were not improved by CPAP, by comparison with sham CPAP. Placebo effects were observed in many of the measures of quality of life and mood (see Table 2.5). Dose-dependent effects (where $p<0.10$) were observed in some outcome measures, indicating that perhaps high compliance was beneficial. However on closer inspection, none of these (apart from the ModMWT and the Epworth Scale) followed a pattern consistent with a physiological effect that was larger than a placebo effect (see Figure 2.11, Figure 2.12, and Figure 2.13).

Figure 2.11: Interactions Between Treatment and Compliance in the Functional Outcomes of Sleepiness Questionnaire- Vigilance Subscale

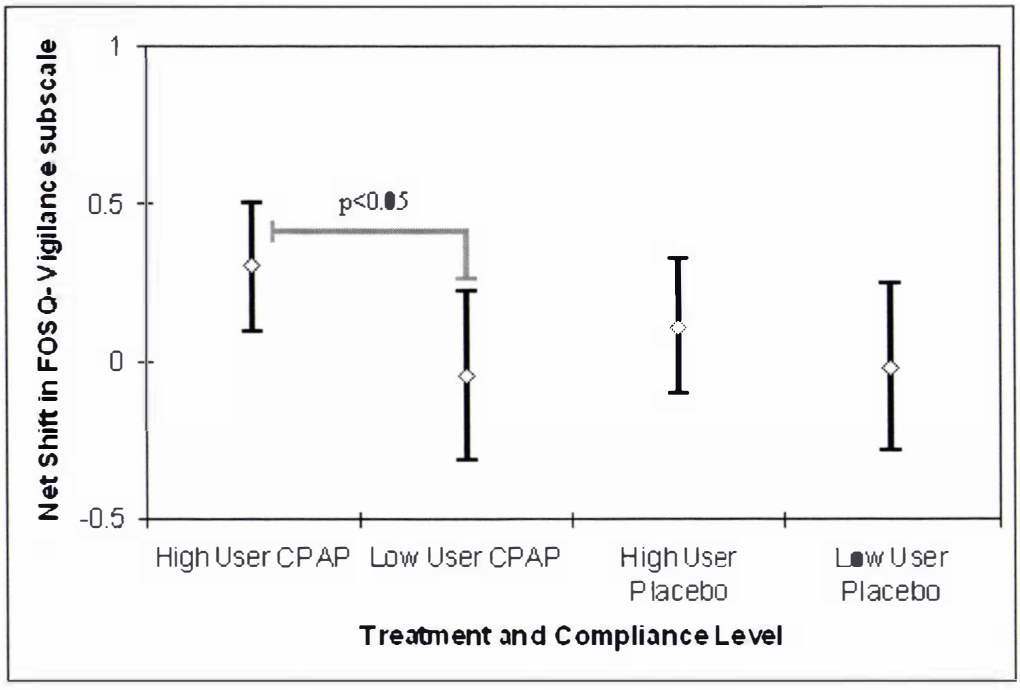


Figure 2.12: Interactions Between Treatment and Compliance in the SF-36 Physical Functioning Subscale.

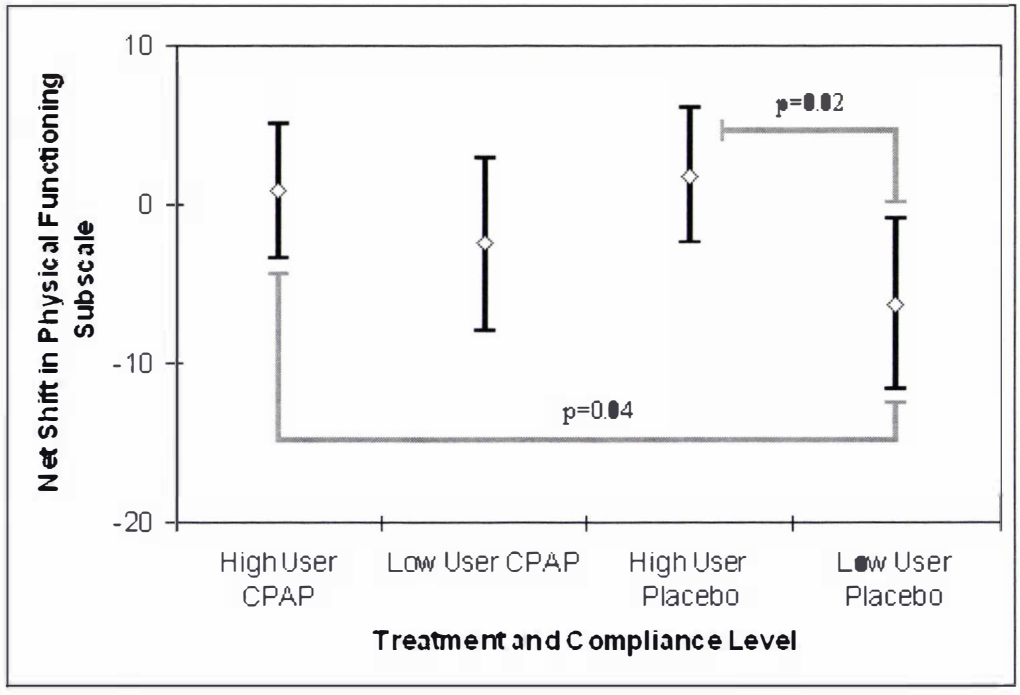
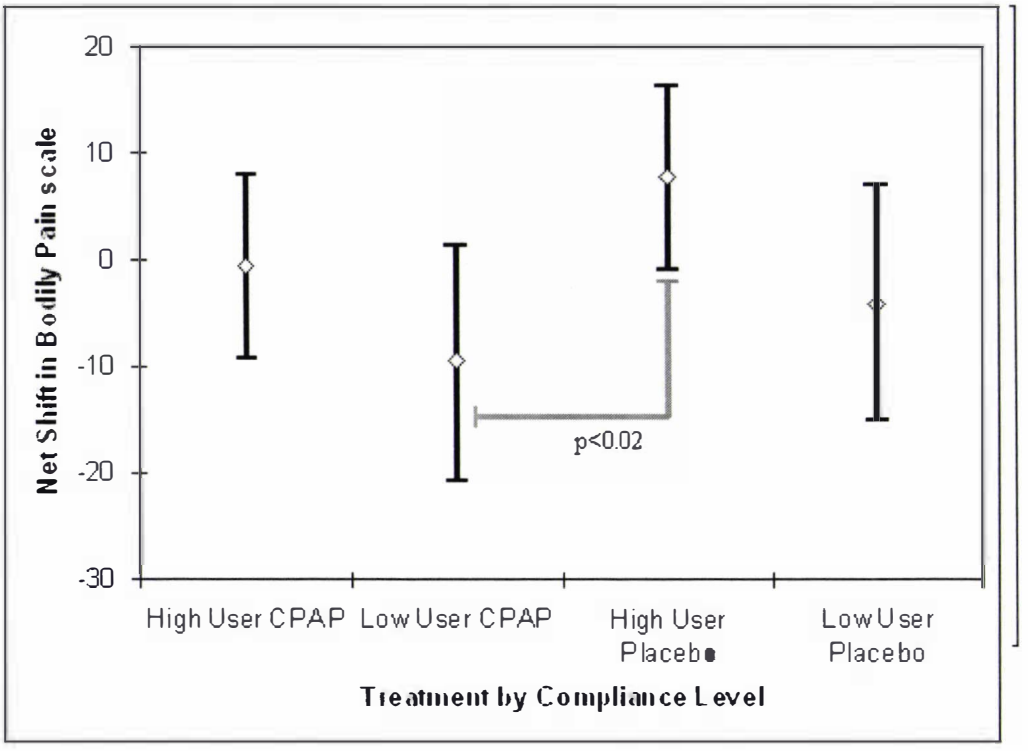


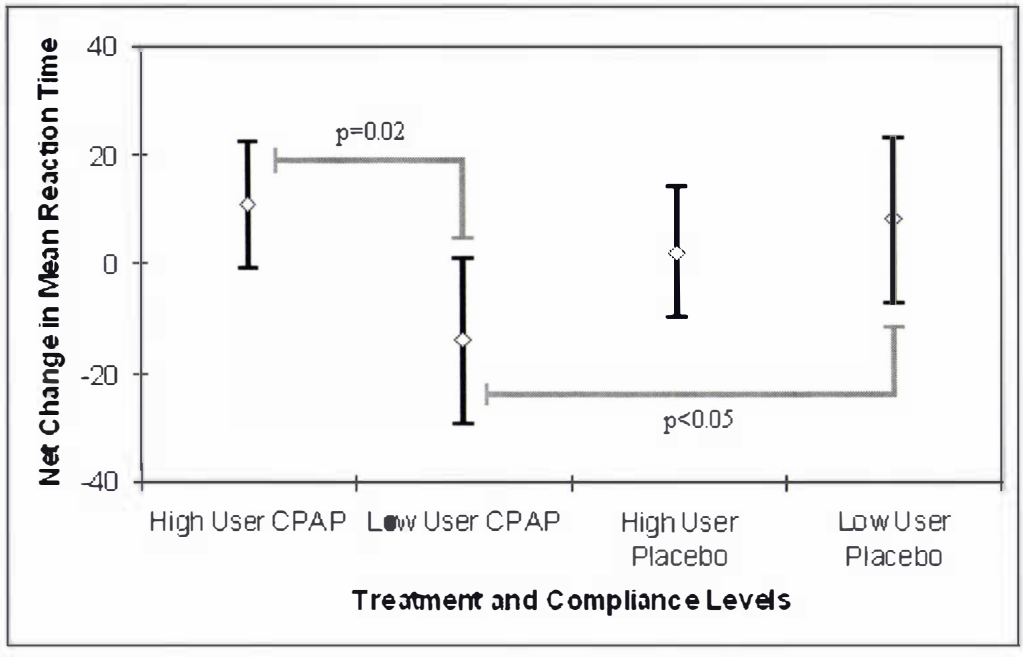
Figure 2.13: Interactions Between Treatment and Compliance in the SF36 Bodily Pain Subscale



In the PVT mean reaction times, a compliance by treatment interaction was found (p=0.03). It was due to reaction times in poor compliers on CPAP treatment tending to worsen (estimated worsening 14ms 95% CI= -30 to 1ms, p=0.067) whilst High users on CPAP (estimated improvement=11ms 95% CI= -1 to 22ms) and Low users of Placebo (estimated improvement=8ms 95% CI= -7 to 23ms) tended to get better (see Figure 2.14). An order effect was also observed in the PVT mean latencies indicating that during the first arm of the trial, latencies were estimated to improve by 10 ms (95% CI= 0.5 to 20ms) whilst worsening by an estimated 7ms in the second arm (95% CI= -16 to 3ms, p for diff= 0.014). Given that this effect is averaged across both treatments, it might show that patients were becoming bored with the measure and had reduced their effort on the task. These are isolated findings and could be due to false positives inherent in the 0.05 threshold, given that 19 mixed models investigated this effect. An order effect was also found in the SF-36 Vigilance subscale, indicating a

estimated improvement over the first arm by 15 points (95% CI= 9 to 22, p<0.001), compared to only 4 points on the second arm (95% CI= -3 to 11, p=0.26, p for difference=0.02).

Figure 2.14: Interactions Between Treatment and Compliance in the Psychomotor Vigilance Task Mean Reaction Time



General patient preference favoured neither treatment with 17/29 preferring placebo (chi-square 0.86, p=0.35). Nineteen thought that placebo treatment was more beneficial for their sleep (chi-square 2.79, p=0.09).

Association Between Baseline Epworth Sleepiness Scores and Treatment Effects.

This trial was too small for many planned subgroup analyses. It is however, possible to look for trends within the entire sample that might indicate which patients are likely to benefit more from CPAP therapy. It is important to limit these multiple investigations to a few useful single measures, due to the expanding possibility of false

positives when undertaking these additional analyses. Epworth Sleepiness Scales are typically administered when patients are referred to sleep clinics. However, there is significant potential for regression to the mean when the Epworth is used both as a screening tool and also as an outcome measure. To minimise the impact of this potential pitfall, each individual's response on placebo measured in this trial was subtracted from the response to CPAP. This should control for both the placebo effect, and possibly regression to the mean. The resulting difference is called the net effect and is plotted below in Figure 2.15.

Figure 2.15: Scatter-gram Showing that Patients Most Affected by Sleepiness at Baseline Benefit Most from CPAP Therapy.

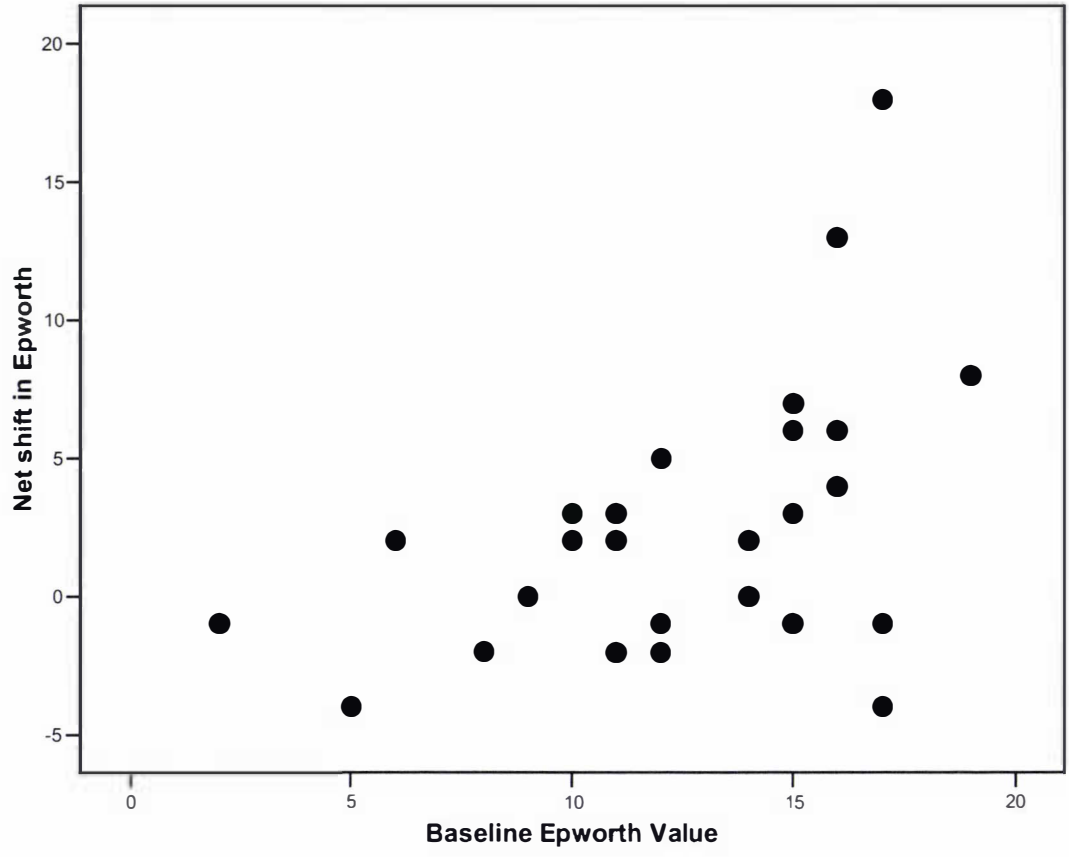
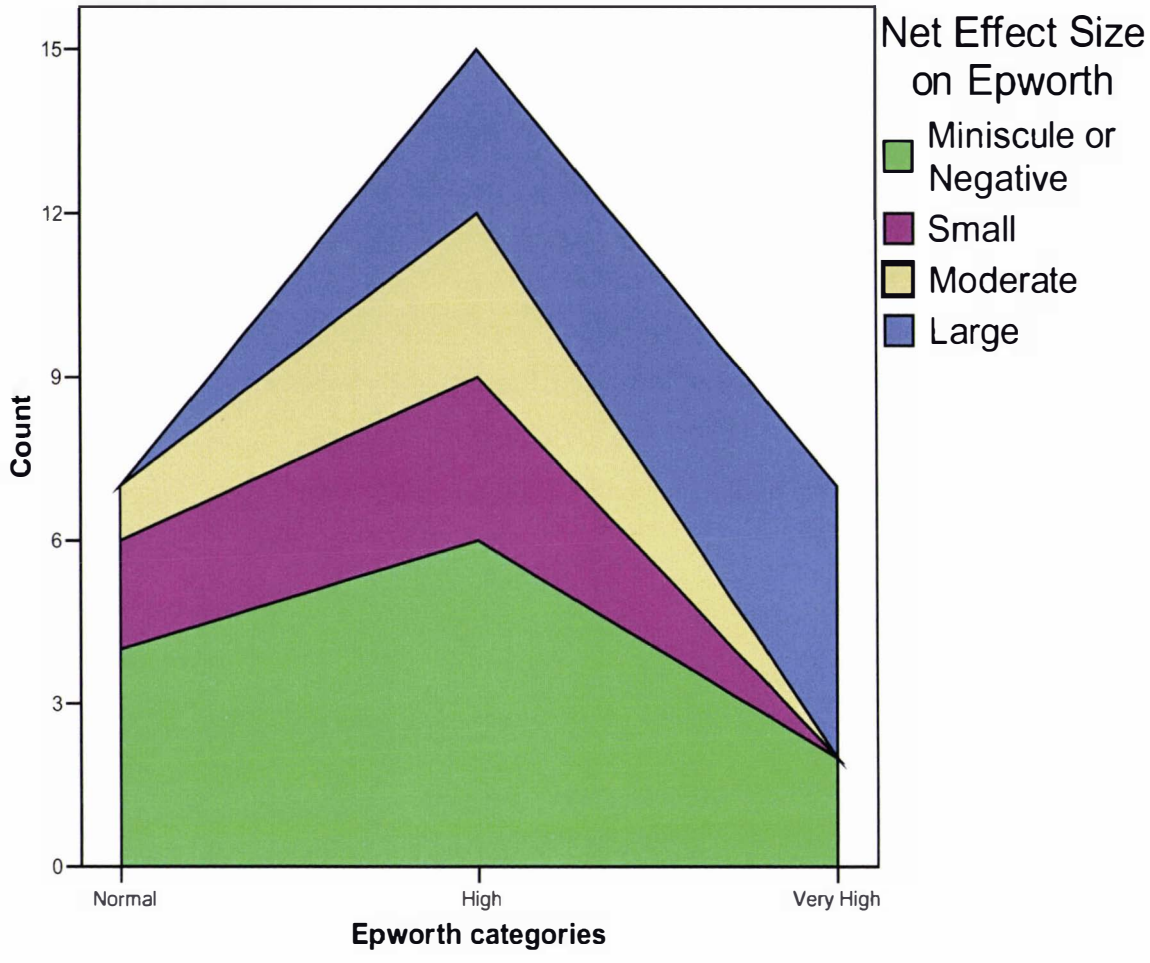


Figure 2.15 shows that each individual's response to CPAP treatment is related to the severity of their daytime sleepiness at baseline, even after controlling for regression to the mean and placebo effects. The relationship between baseline Epworth,

and improvements due to CPAP, is significant (Pearson's r=0.48, p<0.01). The indication is that response to treatment can be predicted by baseline morbidity, although weakly (r2=0.23). This effect has been observed previously and supports Engleman's recommendation that people with mild-moderate SDB (AHI 5-30) may successfully be treated with CPAP, when they have severe sleepiness[210]. Figure 16 shows that those with severe daytime sleepiness (Epworth 16+) are the most likely to gain a large benefit from CPAP therapy, but even in this group, many will not benefit, after adjustment for placebo and regression to the mean effects. Assuming that this sample is representative of the general clinical population, then treating only those with very high Epworth scores (16 or more) will result in about 5/7 patients (71%) having a large clinical improvement. Given that this figure includes an adjustment for (a non-significant) placebo effect, the ratio observed in clinical practice might be slightly higher. If one were to treat only those with high or very high Epworth scores (10 or more) then, based on this data, 8/22 (36%) will have a large response, 3 (14%) a moderate response, 3 (14%) a small response and 8 (36%) will have no response or get worse.

Figure 2.16: Those with Very High Epworth at Baseline are Most Likely to Benefit from CPAP Therapy



Improvements in Objective Wakefulness due to CPAP as a Function of Baseline ESS

Response to treatment, based on Epworth at baseline, can also be assessed by using the Modified MWT. Because the MWT latencies tended to worsen on the placebo arm, and this markedly contributed to the overall effect in the models, the latencies presented here are both the net improvements after placebo adjustment and the smaller individual responses to CPAP alone. Baseline Epworth was only related to the placebo adjusted latencies, and not to the raw CPAP arm latencies. Patients with high, or very high Epworth Scores at baseline were more likely to show a greater improvement in wakefulness on CPAP therapy, after controlling for the placebo effect. However, since

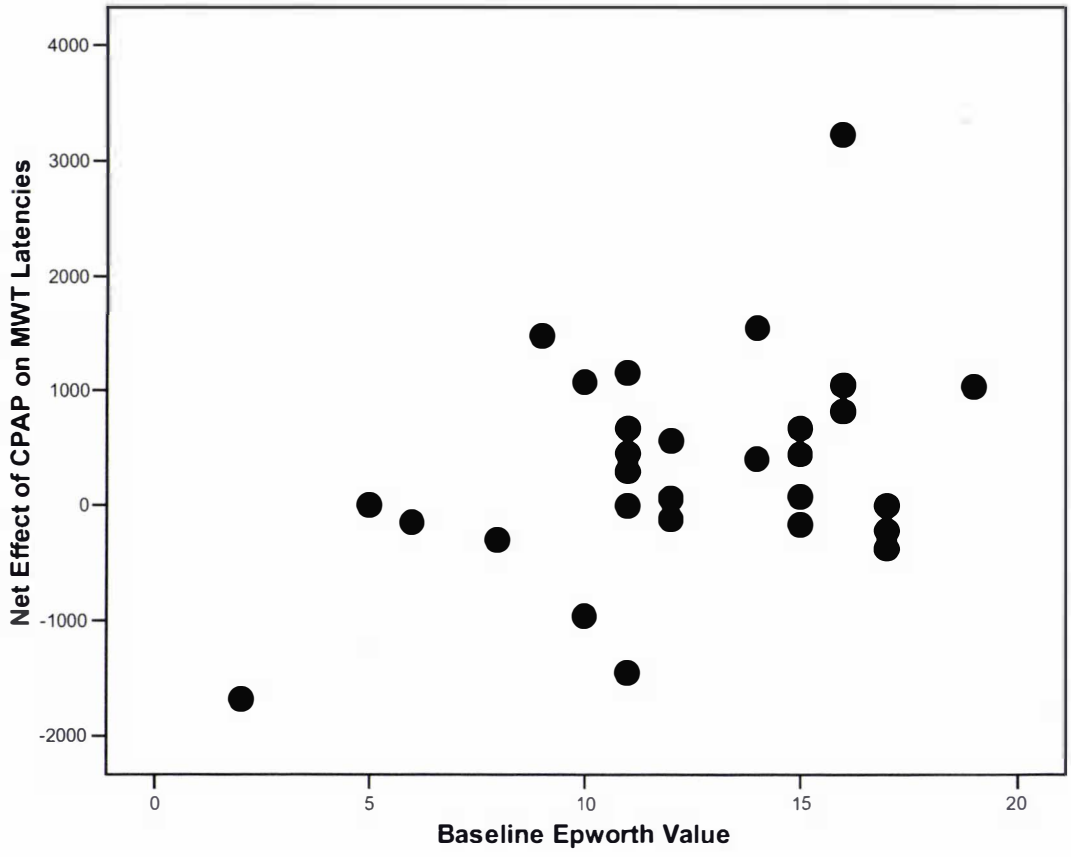
the placebo effect is not quantifiable in clinical practice, baseline Epworth is of no use in a clinical setting for predicting the effectiveness of CPAP for improving wakefulness. Even for the placebo-adjusted relationship, the correlation between baseline ESS and improvement in MWT latencies is very weak (see Table 2.7 and Figure 2.17).

Table 2.7: Correlations Between Baseline Sleepiness Variables and Subsequent Improvements on CPAP

Outcome Measure	Baseline Epworth	Net Effect of CPAP on MWT	Gross Effect of CPAP on MWT
Baseline Epworth Sleepiness Scale	1	**0.394***	0.237
Net Effect of CPAP on MWT		1	**0.687****
Gross Effect of CPAP on MWT			1

Figures are Pearson's Correlation Coefficient (r statistic). Stars represent statistically significant correlations. * significant at the 0.05 level (2-tailed). ** significant at the 0.01 level (2-tailed). Net Effects of CPAP are placebo controlled, gross effects are not placebo controlled.

Figure 2.17: Those with High Epworth at Baseline Have Greater Improvements in Objective Wakefulness.



2.4 DISCUSSION

These findings are somewhat consistent with the overall picture from other similar trials[90, 171, 184-187], and a meta-analysis[149]. For instance, the most recent

study, by Barnes and colleagues[90], shows a significant but small improvement in Epworth attributable to CPAP (1 point), no improvement in MWT latencies (although there was the 2.7 minute significant worsening on the placebo arm, comparable to the magnitude of effect reported here), and an improvement in PVT lapses compared to placebo, but not to baseline. Barnes' FOSQ scores improved by the same amount on CPAP, mandibular advancement splint, or placebo, and in our study CPAP and placebo were not significantly different. The mean SF-36 score in Barnes' 2004 study improved by 4.7 points on CPAP, but only 2 on placebo. Our study showed a clear trend toward placebo outperforming CPAP, with 6 of 8 subscales showing negative net effect sizes. Here, lapses and mean reaction times on the PVT also tended toward favouring placebo over CPAP. Mood, as measured by the Hospital Anxiety and Depression Scale in our study, was also significantly improved under placebo treatment, but not CPAP.

All these placebo effects are somewhat puzzling in the light of a sham CPAP device which is not benign, it worsens sleep quality. Subjective outcomes might be expected to favour placebo, if patients themselves favour the low pressure feel of the interface and expect it to benefit them. Potentially the objective measurements show the true decrements due to placebo (or worsening function due to the continuing effects of OSAS), but are relatively immune to expectation of benefit effects, compared to subjective questionnaires. Other RCTs that used the same outcome measures as this study are tabled below. The patients enrolled in these studies are highly comparable, despite their geographical diversity. Tables further comparing the findings from all similar studies are detailed in Chapter 3.

Qualitatively reviewing the entirety of RCTs addressing the effectiveness of CPAP for treating mild-moderate OSAS gives no clear indication of which, if any, benefits accrue from CPAP therapy.

Table 2.8: Effects on Epworth from Placebo and CPAP: Data From Similar Randomised Controlled Trials

First Author and year	Placebo condition improvement (points of Epworth)	Additional benefit from CPAP (points of Epworth)
Engleman 1997	**4***	0
Engleman 1999	2†	**3** **
Redline 1998	0.4‡	1.1
Monasterio 2001	1.4‡	1.2
Barnes 2002	**2.1***	0.6
Barnes 2004	0.5	**1****
Marshall 2005	0.4	**2.4****

*Placebo effect < 0.05.
** CPAP effect over placebo significant p<0.05
† It is unclear whether this effect was significant or not.
‡ Not a true placebo condition as the alternative treatment is believed to be an active treatment -examples include conservative treatment and positional therapy.

Table 2.9: Effects on Maintenance of Wakefulness Test Sleep Latencies from Placebo and CPAP: Data from Similar Randomised Controlled Trials

Maintenance of Wakefulness Test	Placebo condition improvement (Minutes)	Additional benefit from CPAP (Minutes)
Engleman 1999	No Baseline Taken	1.8
Barnes 2004	**-2.7***	-0.7
Marshall 2005	-3.0	2.2

* Improvement from baseline p<0.05. Negative values from Barnes 2004 and Marshall 2005 indicate that patients sleep latencies worsened whilst on that arm

Table 2.10: Effects on Hospital Anxiety and Depression Scales from Placebo and CPAP: Data from Similar Randomised Controlled Trials

HADS Subscale (each /21)	Anxiety	Depression
Engleman 1997	0.6†	**1.6†***
Engleman 1999	0.9	**1.7‡***
Marshall 2005	-0.8§	-0.3§

*CPAP better than placebo p<0.05
†Placebo effect was not ascertained as baseline values were not taken
‡ It is unclear whether there is a significant placebo effect but it seems probable as the placebo shift is the same as the CPAP only effect, which is significant.
§ Placebo arm was significantly improved from baseline

Table 2.11: Effects of CPAP and Placebo on SF-36 Subscales: Data from Similar Randomised Controlled Trials

SF-36 Scale (/100)	PF	RP	BP	GH	VT	SF	RE	MH	Mean
Redline 1998	-	-	-	-	**8***	-	-	-	-
Engleman 1999	1	17*	7*	2	12†*	11†*	-4†	4†	-
Barnes 2002	-1.3†	0.4	-3.7	-0.6	-0.2†	1.1	-14.2†	0.1†	-
Barnes 2004	-	-	-	-	-	-	-	-	2.7*
Marshall 2005	1	-16†	-7	0	-1†	-8	-12	-3†	-

Figures indicate the amount that CPAP outperformed placebo/control.
*CPAP better than control p<0.05
†Placebo/control arm improved from baseline p<0.05 level.
NS=Not Significant; PC=Physical component summary, MC= Mental component summary, PF=Physical functioning, RP= Role Physical, BP= Bodily Pain, GH= General Health, VT=Vitality, SF= Social functioning, RE= Role emotional, MH= Mental health, Mean= Mean of SF-36 subscales .

Table 2.12: Effects of CPAP on Functional Outcomes of Sleepiness Questionnaire Scales: Data from Similar Randomised Controlled Trials

FOSQ Subscales (/4)	Gen.	Soc.	Act.	Vig.	ISR	Total/20
Monasterio 2001	-	-	-	-	-	NS
Barnes 2002	0.0	0.1†	0.1†	0.1†	0.3	0.2
Barnes 2004	-	-	-	-	-	0.0†
Marshall 2005	0	0.2†	0.1†	0.1	-	0.4

Figures indicate the amount that CPAP outperformed placebo/control.
* CPAP better than control p<0.05
†Placebo/control improved from baseline p<0.05
NS- Not significant and value not reported directly.
Gen.- General Productivity; Soc.- Social Outcomes; Act.- Activity Level; Vig.- Vigilance; ISR- Intimacy and sexual relations.

Table 2.13: Compliance with CPAP Therapy in Similar RCTs

First Author and year	Compliance (Mean hours per night)
Engleman 1997	2.8
Engleman 1999	3.1
Redline 1998	2.8
Monasterio 2001	4.8
Barnes 2002	3.5
Barnes 2004	3.6
Marshall 2005	4.9

A New Meta-Analysis is Now Required

The last published meta-analyses based on a systematic review of the literature included sub-analyses specifically addressing the effectiveness of CPAP for reducing subjective sleepiness in patients with mild-moderate OSAS[149]. The authors concluded that there was insufficient evidence to show that CPAP was effective. The review included literature up to October 2001. Barnes and colleagues published their first study almost a year later, their second in 2004, and the present study was published in 2005[38]. All three new studies have used the Epworth Sleepiness Scale and the Maintenance of Wakefulness Test or the Multiple Sleep Latency Test. Thus, an updated systematic review and meta-analysis was undertaken including only studies of patients classed as having mild-moderate OSAS (AHI 3-30). Such analyses are required because a qualitative overview of the literature does not give a clear indication of whether sleepiness is improved in these patients.

2.5 CONCLUSIONS

This study has shown that humidified CPAP improves subjective sleepiness, and possibly objective wakefulness, in a sample of patients with mild-moderate OSAS. The observed benefit to subjective sleepiness could be viewed as a result of statistical chance, but the pattern of improvement is consistent with an hypothesised physiological dose-response relationship. CPAP did not improve mood, quality of life or psychomotor function.

CHAPTER 3
META ANALYSIS OF RANDOMISED CONTROLLED TRIALS OF CPAP FOR TREATING DAYTIME SLEEPINESS IN PATIENTS WITH MILD-MODERATE OBSTRUCTIVE SLEEP APNOEA SYNDROME

3.1 INTRODUCTION

Systematic Review and Meta-Analysis

A meta-analysis is the statistical combination of more than one study to calculate an average effect that is more accurate than any of the individual studies that it contains. It averages the effects of many studies by giving the most weight to studies that are the most precise (i.e. those with the smallest standard errors of the mean; these studies usually have the largest sample sizes). Meta-analyses are generally of two types; meta-analyses of epidemiological studies aimed at estimating the true size of risk factors for disease (or any other outcome); and meta-analyses of randomised controlled trials aimed at estimating the size of treatment effects. This chapter describes a meta-analysis of the second type.

In order to be classed as the highest standard of EBM, meta-analyses need to be built from all the studies identified and refined by a systematic review[211]. The first stage of this review process is done electronically by searching databases of peer-reviewed published studies such as MEDLINE. Database searching for content for systematic reviews can normally be achieved by the combination of Medical Subject Headings (MeSH) into and/or instructions in MEDLINE. These searches give highly replicable results as long as the same dates are used. The second stage is potentially less reliable and involves the inclusion or exclusion of all identified studies by the rigorous

application of predetermined criteria for appropriateness. These criteria, when properly designed, should leave a group of studies that are all directly applicable to the research question. They should also retain all such applicable studies. If another group of reviewers were to repeat the process, they should end with the same set of studies.

A meta-analysis based on a systematic search of all available literature aims to reduce the sources of bias that have resulted in qualitative reviews reaching different conclusions regarding the same question. Qualitative reviewers can choose to ignore, or promote certain studies, that fit their own particular opinions on a subject. Placing undue weight on favoured studies, and ignoring or downplaying methodologically sound studies that have differing results, can result in differing conclusions.

The combination of systematic review and meta-analysis should result in the highest quality evidence for treatment effects. It stands above the level of proof offered by a single RCT, or a small number of RCTs, that have been meta-analysed (it is possible to meta-analyse as few as two trials and to do so without systematic review). However, sometimes meta-analyses of small RCTs (such as the one reported here), can have findings that are later overturned by the findings of a mega-trial (a very large RCT). Ongoing studies likely to affect the findings presented here are discussed later.

Why use this technique here?

Firstly the published trials are from disparate geographical locations (Spain, Scotland, USA, and Australia). This situation might encourage regionalism, where the findings of the local RCT(s) are highlighted, and other findings ignored or marginalised

in qualitative reviews or local guidelines. Secondly, all of the trials are small by the standards of most of scientific medicine (<150 participants) and have limited statistical power to detect subtle treatment effects. A systematic review and meta-analysis might help solve these problems by succinctly summarizing the mean effect of all published studies.

Whether to use CPAP to treat sleepiness in people with mild-moderate OSAS is still an open question. The previous systematic review and meta-analysis found insufficient evidence to conclude that CPAP (the gold standard for treatment for severe OSAS) reduced subjective sleepiness in patients with mild-moderate OSAS[149].

Previous Meta-Analyses of CPAP for all severities of OSAS.

Wright et al. (1997)

The first systematic review and meta-analyses (see also Chapter 1), by Wright and colleagues[12], attempted to answer a number of broad questions, including whether CPAP or any other method effectively treated OSAS. Only one published RCT was found[142], but the authors chose to ignore it when drawing their conclusions. The paucity of RCT-based evidence was interpreted as a lack of CPAP efficacy and public funding for CPAP therapy in the United Kingdom was withdrawn as a result. In 2000 an updated Cochrane review of CPAP efficacy noted that sufficient evidence was now available to show that severe OSAS could be successfully treated with CPAP, but that in the mild-moderate severity range, further trials were required[212].

The Australasian Meta-Analyses (1996, 2000)

The Australasian Health Technology Advisory Committee (AHTAC), in conjunction with the New Zealand Ministry of Health, released a very unusual draft meta-analysis in 1996[148] that attempted to deal with the same questions as Wright et al.[12]. It faced the same lack of evidence, but came to completely different conclusions. They chose an unorthodox method to deal with the lack of RCT-based evidence by choosing to meta-analyse uncontrolled studies. This, despite the considerable possible influence of placebo effects, which were unfortunately not quantified by Engleman and colleagues in the only extant placebo controlled study[142]. The AHTAC meta-analysis tended to focus on overnight polysomnographic indices of SDB as a treatment outcome in OSAS. From the single RCT (mean AHI=28), and a much larger quantity of lower level evidence, they concluded that the effects of CPAP on OSAS were large and found across a number of different outcomes. It was recommended that patients who should be treated were those with:

a. AHI above 30 and symptomatic OSAS (at least 2 of the following)
a. Chronic or persistent snoring, choking or observed apnoeas
b. Daytime sleepiness
c. Disturbance of concentration, memory or mood

b. AHI 20-30 and one of the following;
a. Pathological Sleepiness
b. Respiratory Failure
c. Symptomatic cardiovascular or cerebrovascular disease
d. Severe Desaturation ($SaO2 \leq 75\%$)

Whilst the (a) category of patients was evidence-based advice, and backed by the single RCT, the second was probably not. The mean AHI in the uncontrolled studies

was calculated to be about 60/hr, and did not deal with patients in the AHI range 20-30. Whilst the second category sounded like sensible advice, it seemed to be the result of clinical consensus building. It should not have been listed as a recommendation arising from a meta-analysis. Firstly it does not match the evidence meta-analysed. Secondly, the meta-analysis was not built with randomised controlled trials, and thus the raw data cannot be relied upon.

Despite it's shortcomings, this report seems to have been widely accepted in Australasia and was used as a treatment guideline in the Wellington clinic where the studies from Chapters 2 and 4 were carried out. The further recommendations of the report listed the major weaknesses in sleep medicine evidence that required further research. Some of this much sought after information is now available from such sources as the ongoing Sleep Heart Health, and Wisconsin cohorts and the much improved situation regarding RCT-based treatment evidence. However, gaps in the evidence remain. For instance there is still no consensus as to which neuropsychological batteries/tests should be employed in OSAS research, making meta-analyses of changes in neuropsychological function extremely difficult.

The final version of the meta-analysis was eventually released in 2000 by the Australian National Health and Medical Research Council[111]. By this stage (search in 1999), twelve RCTs had been published and the unorthodox approach of combining uncontrolled studies could be abandoned. The 2000 report found a significant reduction in mean Epworth Scores of 1.1 points (95% CI= 0.32 to 1.88) across 6 studies involving patients across a broad spectrum of OSAS severity. Four of these studies were in the mild-moderate range[171, 184, 185, 187]. Once these were removed, the two remaining severe studies[161, 213] indicated an average reduction of about 5.8 points (95% CI 3.4

to 8.3 points). A further RCT reported a median reduction of about 6.5 points[150], but unfortunately could not be included in the analysis because the standard error could not be calculated. Nevertheless, it did help support the estimated magnitude of the effect. Four studies had used the MSLT as an outcome, but meta-analysis revealed no significant benefit from CPAP therapy (mean improvement 0.35 mins, 95% CI -0.22 to 0.91). However, the one study investigating those with severe OSAS did show a benefit of about 2.4 minutes[151]. Various other meta-analysed neuropsychological outcomes indicated benefits to Trailmaking Test B and PASAT 2 second test, but not to Steer Clear or the Digit Symbol Substitution test. The authors also found that general health-related quality of life, energy and vitality, all improved. Depression measures also improved in patients with both severe and mild OSAS. Once again, the recommendations for treatment were probably seen as sensible, but did not match the evidence that had jut been identified and combined.

The recommendations were as follows:

a. AHl>20 and symptomatic during the day with at least two of:
a. Chronic or persistent snoring, choking or observed apnoeas
b. Daytime sleepiness
c. Disturbance of concentration, memory or mood
b. AHI 10-20 plus two or more of the above conditions and at least one of the following:
a. Pathological Sleepiness
b. Respiratory Failure
c. Symptomatic cardiovascular or cerebrovascular disease
d. Severe Desaturation (SaO2 $\leq$ 75%)

None of the new trials reviewed indicated that treating people with symptomatic mild OSAS in the range of 10-20 AHI was effective. Furthermore none of the trials have ever used AHI>20 as an entry criteria. Trials have used AHI 5-15, AHI 5-30,

AHI>5, AHI>10, and AHI>30, as criteria (sometimes these have been de-saturation indices rather than AHI). The severe syndrome trials tend to be AHI>30, but the bulk of trials are actually in the mild-moderate range because that is where much of the clinical uncertainty about CPAP remains. Thus an evidence-based recommendation should probably be built around the most commonly used and recommended severity thresholds. These are mild 5-15, moderate 15-30 and severe 30+[14]. Randomised controlled trial evidence exists at these cut points, and yet the recommendations ignore them and build additional criteria that were not tested in the trials reviewed. So whilst some evidence does exist to indicate CPAP might be effective in mild (AHI 5-15) or mild-moderate (AHI 5-30), none of the RCT studies have used cutpoints at 10-20, or at 20 and above. It is unclear why these particular thresholds were used by the NHMRC (2000). In 1996 the recommendations might have been defensible of the basis of best guess of clinical consensus in the environment of almost no evidence, but by 2000 much more evidence was available. In fact the 1996 recommendations were better supported by the evidence base available in 2000 than were the 2000 recommendations.

Although it was produced well after the NHMRC report, our own RCT (Chapter 2) illustrates this lack of match between recommendations and evidence. The inclusion criterion was AHI 5-30/hr, and our mean AHI was about 22. This means that most of our patients had a diagnostic AHI of 20-30/hr. According to the NHMRC 2000 recommendations, these patients should be treated with CPAP, and yet our patients showed no significant benefits from treatment other than a moderate reduction in the Epworth Sleepiness Scale[38].

Patel et al. (2003)

The third and most recent meta-analysis answered a very focussed question, compared to previous studies[149]. It looked only at the effects of CPAP on the three sleepiness measures which are the most well-defined and widely used in patients with OSAS; the Epworth Sleepiness Scale, and the EEG based Maintenance of Wakefulness Test (MWT) and Multiple Sleep Latency Tests (MSLT). It also accepted the use of the non-EEG based OSLER[158] test of maintenance of attention, as a homologue of the MWT.

Patel et al. found significant benefits for CPAP over placebo in reducing both subjective and objective daytime sleepiness in people with a wide severity range of OSAS[149]. Epworth Scores dropped by a mean 2.94 points (95% CI 1.61 to 4.26). A pooled metric of the MWT and MSLT, expressed as a net shift in minutes of sleep latency, improved by a mean 0.93 minutes (95% CI 0.1 to 1.76 mins). In patients with severe OSAS (AHI tended to be >30), Epworth scores improved by 4.75 points (95% CI 2.97 to 6.53).

However, most people with OSAS have mild to moderate severity[3, 14]. Despite the 4 published randomised controlled trials found[184-187], there was insufficient evidence to conclude that CPAP reduced sleepiness in these patients. The mean Epworth reduction was 1.1 points (95% CI -0.13 to 2.32). However, Patel et al. included one additional study that specifically investigated the use of CPAP in participants with AHI>30, but no daytime sleepiness[83]. This study should not have been included as it addressed an important, but separate, question regarding treatment of severe but asymptomatic SDB.

Patel and colleagues also combined the metrics of the MWT and MSLT into a single measure- shift in daytime sleep latency. This is an unusual approach, as it is generally felt that the two tests measure different abilities or characteristics[152, 153, 214, 215]. Sangal and colleagues, in 1992, reported that the MSLT and MWT are only weakly correlated (r=0.41), and that the MSLT and MWT are differentially sensitive to the effects of CPAP treatment[152, 153]. However, Patel and colleagues, found that the mean reduction in MSLT (0.74 mins) was not significantly different from the reduction in MWT (3.0 mins, p for diff=0.12). They also combined two different versions of the MWT, the standard EEG based measure[194, 195] and a behavioural test of alertness used in Oxford called the OSLER[158]. These two tests might be differentially sensitive to CPAP treatment.

Since the time of that literature search, at least another three studies in the mild to moderate range have been published[38, 90, 171]. These three studies are all crossover studies and contain 137 patients in addition to the 272 patients in the four studies in the previous systematic review(two crossover, two parallel, not including the 54 patients treated by Barbe et al.[83]). Given Patel's earlier lack of statistical power in the mild-moderate range, the inclusion of a misclassified study, and the combination of two usually incompatible measures, it seems timely to repeat the systematic search and meta-analysis with the specific intention of answering whether CPAP reduces sleepiness in patients with mild-moderate OSAS. Three groups of analyses summarising the effects of CPAP on sleepiness are required. These are the effects of CPAP on the Epworth Score, the Maintenance of Wakefulness Test, and the Multiple Sleep Latency Test. The Epworth score is usually measured with either the Maintenance of Wakefulness Test or the Multiple Sleep Latency Test. The MSLT and MWT are only very rarely measured together, given their onerous and time consuming nature[152].

Unfortunately this will mean that the MWT and MSLT analyses have less statistical power than the Epworth Analyses. A review of measures can be found in Appendix 1.

3.2 METHODS

Systematic Review of Randomised Controlled Trials

Systematic Search

A systematic search was undertaken using MEDLINE listed publications from 1st January 1994 until 31st December 2004. The restricted search dates were used because no previous systematic search has identified a suitable study before 1994. Keywords used to identify randomised placebo controlled trials of CPAP in the treatment of mild-moderate OSAS were (apnoea.af or apnea.af or hypopnoea.af or hypopnea.af) AND (CPAP.af or continuous positive airway pressure.af or positive airways pressure.af or positive pressure.af) AND (randomized controlled trial.pt or .ti or clinical trial.pt or .ti). Suffixes indicated search fields used in MEDLINE (.af=all fields, .pt=publication type, .ti=title). These search criteria were closely modelled on those used by Patel[149]. The findings from Chapter 2[38] were included as being found from a hand search. The bibliographies of previous meta-analyses and RCTs found from the systematic search were also hand-searched for additional studies.

Systematic Review of Identified Studies

The systematic review was independently conducted by myself and Dr Maree Barnes of the Institute of Breathing and Sleep at the Austin Hospital in Melbourne Australia

Study abstracts, and if necessary, full paper versions of studies, were used to review and include or exclude studies according to the following criteria; only studies that were randomised controlled trials with treatment arms of at least one week of manually titrated CPAP therapy undertaken on adults with primarily obstructive sleep apnoea syndrome were considered. Multi-comorbid studies, where combinations of disorders such as OSAS and heart failure, or OSAS and gastro oesophageal reflux disease, were excluded.

The control arm or arms had to be either conservative management advice or a placebo. Conservative management/treatment of OSAS usually includes advice about weight loss, sleep hygiene, avoidance of supine sleep posture, and avoidance of alcohol and sedatives. Conservative management is not technically a placebo, but has yet to be shown to be an effective treatment over the short term. Weight loss is probably beneficial over a longer term, given cohort findings[52], but this has not been tested[139]. Acceptable placebos included orally ingested placebo tablets (pills) or sham CPAP devices that did not improve overnight indices of SDB compared to polysomnography at baseline[169]. Unacceptable comparisons to CPAP included any device that improved SDB, such as suboptimally pressured but partially therapeutic CPAP, or any surgical techniques (including sham surgery), or sham or ineffective mandibular advancement splints. Studies not exclusively investigating mild-moderate syndrome classification (apnea hypopnea index specified as 5-30/hour for all patients) were excluded. Trials had to measure the Epworth Sleepiness Scale for chronic subjective sleepiness[193], and/or the Multiple Sleep Latency Test of objective sleepiness[216], and/or the Maintenance of Wakefulness Test of objective wakefulness[194] in order to be analysed.

Meta-Analyses of Randomised Controlled Trials

Data Extraction and Trial Quality Assessment

Study characteristics retrieved for each trial report included the number of patients completing the entire study, placebo type, trial structure, length of follow-up, gender balance, average age, average Body Mass Index (kg/m2), average AHI, method of measuring airflow (thermistor or pressure transducer), dropout rate, average CPAP use per night, country of origin, single venue vs. multi-site, and baseline severity of sleepiness, as measured by the 3 potential outcome measures. The primary data analyses required the mean and standard error of the mean, for each treatment response to CPAP, after placebo effects were controlled. Jadad scores of trial quality were calculated for each study. The Jadad score[217] is a validated metric used in the quality assessment of randomised controlled trials. Points are awarded between 0 and 5 for appropriate double blinding (2 points), appropriate randomisation (2 points), and the clear reporting of trial dropouts and exclusions (1 point). Jadad scores can be used in meta-analyses to include and exclude studies, and to assess whether trial quality has affected the size of the treatment effect, and thus could explain heterogeneity of treatment effects between studies.

Data Synthesis and Statistical Analysis

Authors of individual studies were contacted to provide the differences in treatment effect between CPAP and comparison, and standard errors where these were not available from the published paper. Authors were also contacted if any of the information provided in Table 3.2 and Table 3.3was absent from the published version of the trials.

Each net mean treatment effect, and the standard error of that mean effect, were combined to produce a pooled estimates of treatment response, with both fixed and random effects model assumptions (META command, v7.0 Stata Corp , College Station, TX). Random effect models assume that some of the difference in the size of effects is due to different sample origins, resulting in real differences in true treatment effect. Fixed effects models assume that studies are generally comparable and are converging on the same true effect, and observed differences are assumed to be artefacts of random sampling.

Heterogeneity in treatment responses to CPAP, after placebo adjustment, was tested between studies with the Q statistic. Where heterogeneity existed meta-regressions were performed to explore the possible sources of heterogeneity (MetREG command produces Q statistic). Variables potentially associated with CPAP efficacy (trial and patient characteristics- found in Table 3.2 and Table 3.3) were successively incorporated into a single covariate model to assess their contribution to heterogeneity.

Each study was excluded sequentially, to measure the influence of individual study esimates on the pooled estimate (METAINF command). Publication bias was evaluated visually using the funnel plot, and statistically (METABIAS command) using Egger's and Begg's tests. Meta-analyses and sub-analyses were carried out by Noemi Travier (Sleep/Wake Research Centre Biostatistician) using my suggestions for testing, and were interpreted by both of us.

Effect sizes[112] for significant treatment effects were calculated by dividing the mean size of the effect identified in the meta-analysis by the baseline standard deviation of a suitable reference group of 110 Australian mild-moderate OSAS

patients[218]. Most of these patients were also involved in one of the Australian mild-moderate RCTs[90], and are thus an ideal reference group from which to derive reference standard deviations.

3.3 RESULTS

Systematic Review

The systematic search of MEDLINE produced 295 study abstracts of further interest. A further suitable study (Chapter 2) was also included[38]. The search terms used were not specific to our question, but should have been sensitive. Consequently the majority of the studies identified were not relevant to the research question. Many identified studies were not randomised controlled trials, did not investigate OSAS, or involved paediatric patients, or patients with OSAS and other serious comorbidities (such as diabetes or heart failure). Many identified RCTs that investigated CPAP (or a similar such device) compared to either mandibular advancement splints, auto titrating CPAP, bilevel PAP or a mixture of surgical procedures. There were also a number of trials that employed what was described as a placebo CPAP machine that substantially reduced AHI, and thus was partially technically effective[82, 157, 162, 163, 219-221].

Another trial compared positional therapy to CPAP in a group of patients who had almost exclusively mild-moderate OSAS[222]. However, this OSAS was mostly positional, and thus the positional therapy did not constitute a placebo in this particular patient group. One further trial came close to inclusion as it included a sub-analysis of specifically mild to moderate OSAS (AHI 10-30) in an RCT[223]. However, the study was aimed at comparing the economic effects of immediate treatment with CPAP, with a six month delay (without diagnostic polysomnography and with conservative

management), and it was unclear whether patients were given the impression that any equipoise existed between conservative management advice and CPAP. If patients were led to believe that 6 months of conservative management was somehow inferior to CPAP, then the CPAP arm of the trial would contain some amount of uncontrollable placebo effect, making this trial unsuitable for inclusion in the meta-analysis. The six month delay in diagnostic PSG also means that the control and CPAP groups systematically differed in their interactions with investigators. The trial was therefore excluded from the analyses, as a non-placebo controlled study. Ip and colleagues[224] compared the effects of CPAP to observation only. Observation only is not usually regarded as a placebo and the study was thus excluded. Twenty-seven studies remained as suitable randomised placebo controlled trials of CPAP, and were screened for inclusion if they exclusively included patients with AHI 5-30.

Both reviewers agreed that there were seven RCTs investigated patients with exclusively mild-moderate OSAS (AHI 5-30). Brief descriptions of the 21 excluded RCT studies can be found in Table 3.1. Figure 3.1 is the trial inclusion flow chart. All seven mild-moderate RCTs employed the Epworth Sleepiness Scale, four used the Multiple Sleep Latency Test, and the remaining three the Maintenance of Wakefulness Test. Characteristics of included trials can be found in Table 3.2. Table 3.3 contains the summary of patient characteristics from each of the included trials.

Table 3.1: Excluded Placebo Controlled CPAP RCTs

Source First Author, Year	No. of Patients	Control type	Crossover study?	Treatment duration (weeks)	Entry Criteria	MSLT or MWT Measured?	Epworth Measured?
Engleman 1998	23	Pill	Y	4	AHI≥15	MSLT	Y
Engleman 1996	13	Pill	Y	3	AHI≥5	None	N
Lojander 1996	44	CT	N	52	ODI4%≥10	None	N
Engleman 1994	32	Pill	Y	4	AHI≥5	MSLT	N
Meston 2003	101	Sham CPAP	N	4	ODI4%≥10	None	N
Robinson 2003	101	Sham CPAP	N	4	ODI4%≥10	None	N
Chakravorty 2002	71	CT	N	12	AHI≥15	None	Y
Pepperell 2001	95	Sham CPAP	N	4	ODI4%≥10	Osler	Y
McArdle 2001	22	Pill	Y	4	AHI>15	None	Y
Montserrat 2001	45	Sham CPAP	partial	6	AHI≥10	None	Y
McArdle 2001	22	Pill	Y	4	AHI>15	None	Y
Barbe 2001	54	Sham CPAP	N	6	AHI>30	MSLT	Y

Neleson 2001	41	Sham CPAP	N	1	RDI≥15	None	N
Henke 2001	46	Sham CPAP	Partial	2.5	AHI>10	None	Y
Faccenda 2001	68	Pill	Y	4	AHI≥15	None	Y
Stradling 2000	101	Sham CPAP	N	4	ODI4%≥10	Osler	Y
Hack 2000	59	Sham CPAP	N	4	ODI4%≥10	Osler	Y
Jenkinson 1999	101	Sham CPAP	N	4	ODI4%≥10	Osler	Y
Lojander 1999	27	CT	N	52	ODI4%≥10	None	N
Ballester 1999	105	CT	N	12	AHI>15	None	Y
Robinson 2004	220	Sham CPAP	N	4	ODI4%≥10	None	N
Pelletier-Fleury 2004	108	CT	N	26	AHI 10-30	None	Y
Pelletier-Fleury 2004	63	CT	N	26	AHI≥30	None	Y

Pelletier-Fleury 2004[223] is a single study with subanalyses that split mild-moderate OSAS away from severe. This is done for clarification purposes only for this table. The mild-moderate study was not included in analyses because the aim of the study was not to test the effectiveness of CPAP, but to test the effects of six months delay in diagnostic polysomnography. The OSLER is a non-EEG based version of the MWT used in primarily in Oxford[158]

Six included studies scored 3 on the Jadad scale of trial quality (see Table 3.2). All these had adequate descriptions of trial dropouts and withdrawals, in addition to being appropriately randomised. In one study it was unclear for what reasons dropouts had occurred and thus the study scored a 2 on the Jadad scale[186]. No studies were truly double-blinded, as CPAP titrations require that at least some of the patients' interactions with sleep staff require those staff to be unblinded. This is true even when sham CPAP devices are used[38]. This reasoning differs from that of Patel, who thought that studies employing sham CPAP could be effectively double-blinded and were given Jadad scores of 5 where appropriate. Thus, a study such as that reported in Chapter 2 would probably have been given a 5 according to the criteria of Patel et al.[149].

Figure 3.1: Systematic Search Flow Chart

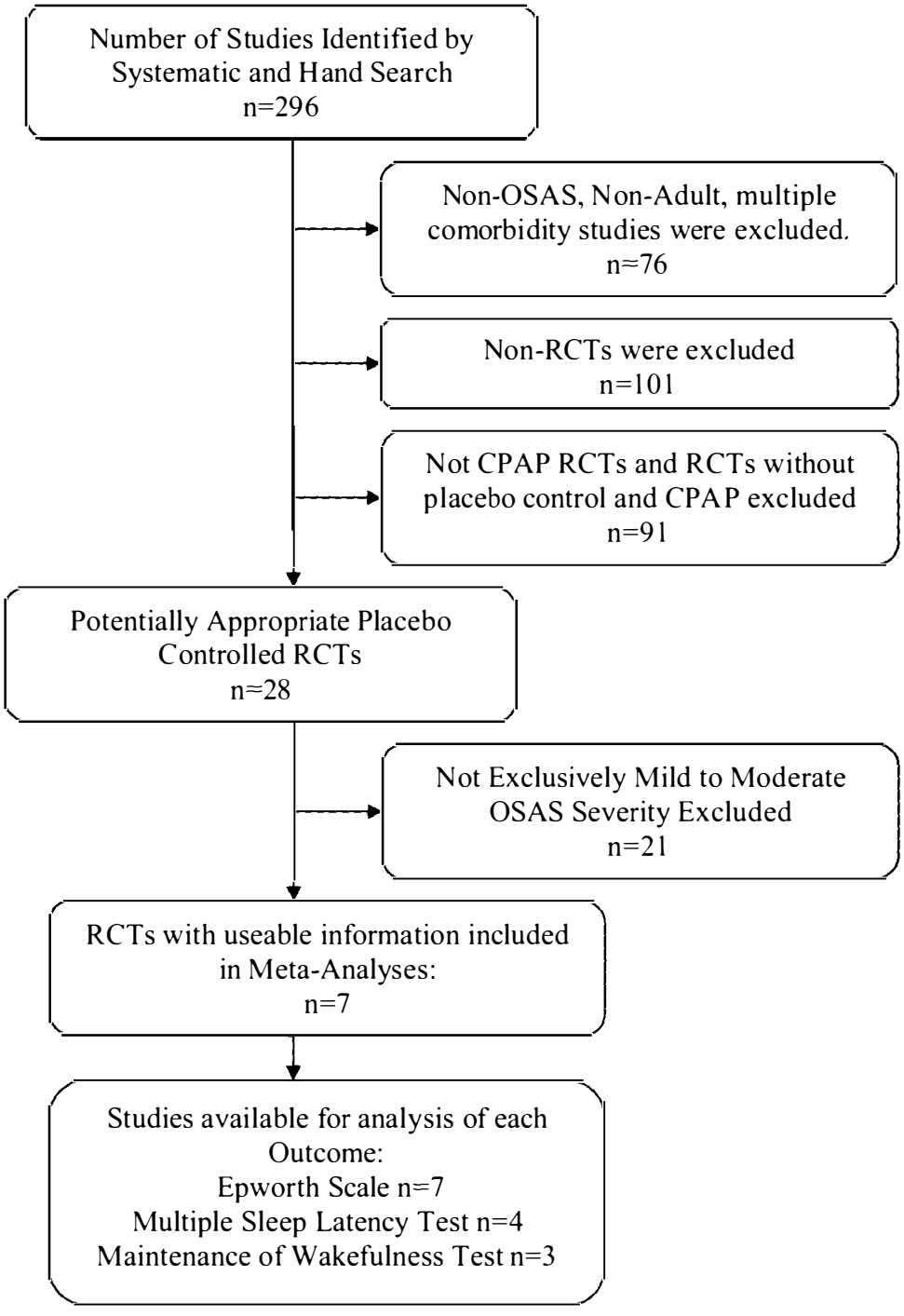


Table 3.2: Study Descriptions

Source Author, year	No. of Patients	Control type	Crossover study	Treatment duration (wks)	Drop Out Rate	MSLT or MWT	Jadad score
Engleman 1997	16 and 9	Pill	Y	4	11%	MSLT	3
Redline 1998	97	CT	N	10.5	13%	MSLT	3
Engleman 1999	34	Pill	Y	4	8%	MWT	3
Monasterio 2001	125	CT	N	24	12%	MSLT	2
Barnes 2002	28	Pill	Y	8	33%	MSLT	3
Barnes 2004	80	Pill	N	12	30%	MWT	3
Marshall 2005	29	Sham CPAP	Y	3	6%	MWT	3

CT= Conservative Treatment

Table 3.3: Patient Characteristics: Analysed Studies

Source Author, Year	Mean Age	Sex Ratio (n males)	Mean Body Mass (Kg/M^2)	CPAP Use hrs/night	Baseline Epworth	Baseline MWT (Mins)	Baseline MSLT (Mins)
Engleman 1997	52	75% (12)	30	2.8	14		10
Redline 1998	48	52% (50)	33	3.1	10		10
Engleman 1999	44	61% (21)	30	2.8	13	NR	
Monasterio 2001	54	86% (108)	29	4.8	12		10.5
Barnes 2002	45	83% (35)	30	3.5	11		12.5
Barnes 2004	46	79% (63)	31	3.6	10.2	30.7	
Marshall 2005	51	76% (22)	31	4.9	12.5	20.9	

NR= No Baseline Measured/ Reported

Meta-Analyses- Main Results

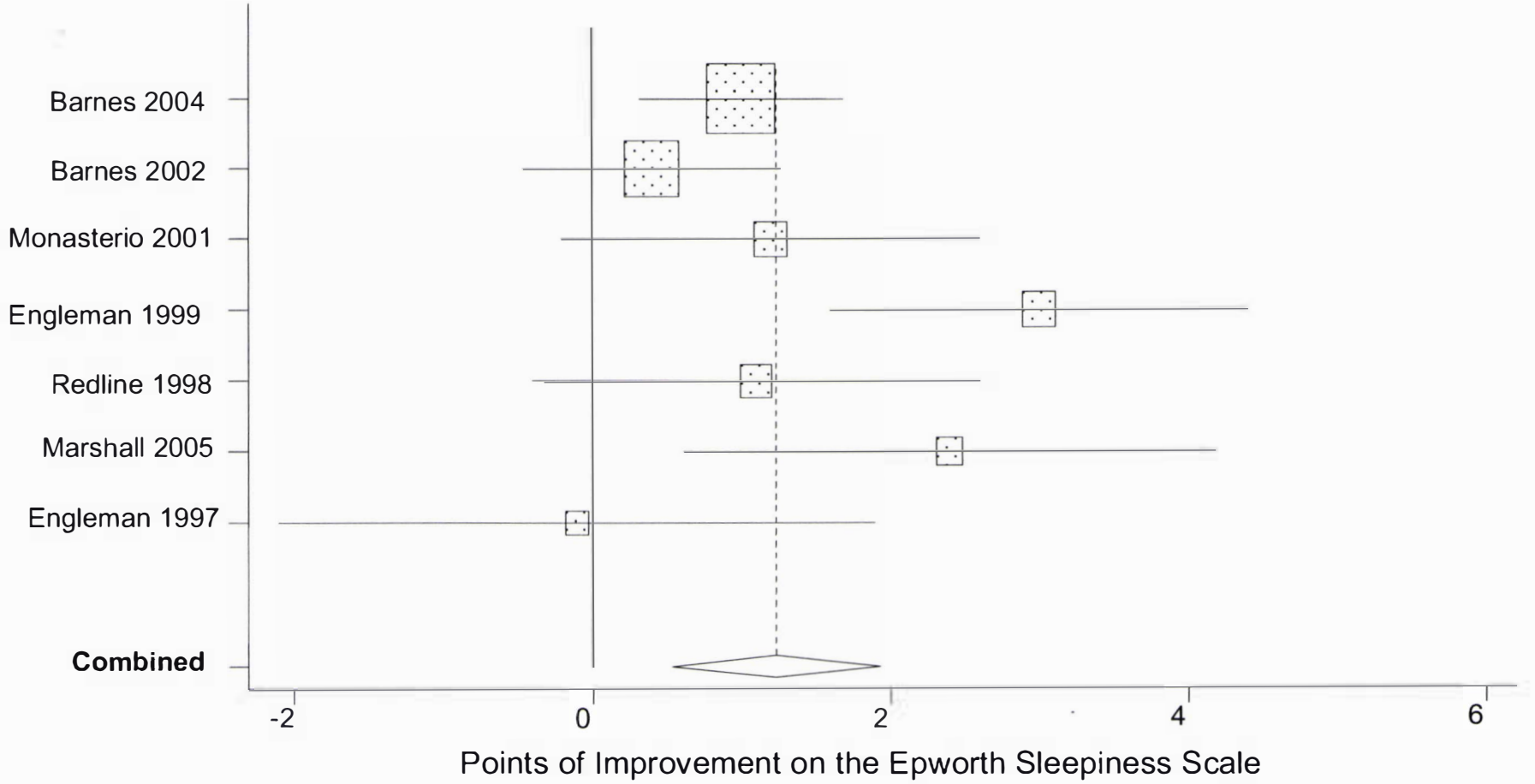
Epworth Sleepiness Scale

Epworth scores were used in all 7 studies (see Table 3.4). Using the random effects method, the Epworth Score was found to improve by 1.2 points (95% CI 0.5 - 1.9, p=0.001, see Figure 3.2). The fixed effects method gave very similar results, with Epworth improving by 1.1 points (95% CI 0.7 – 1.5, p<0.001). Epworth Sleepiness Scale effects were found to be heterogeneous (Q=13.1, df=6, p=0.04). Meta-regression indicated that heterogeneity was not related to the following trial characteristics: Jadad quality score (p=N/A), design of trial (parallel vs. crossover, p=0.88), control comparison (p=0.4), treatment duration (p=0.68), multi vs. single site (p=0.11), or country of origin (all p>0.3). Nor was it due to the following patient characteristics: baseline Epworth scores (p=0.46), age (p=0.69), mean body mass indices (p=0.96), percentage of males in the study (p=0.23), syndrome severity (mean apnoea hypopnoea index, p=0.92), method of measuring airflow (p=0.49), or average nightly compliance with therapy (p=0.77). However, the percentage of dropouts in each study did explain some heterogeneity. Studies that had lower dropout rates had more positive effects upon Epworth (Z=2.38, p=0.02).

Table 3.4: Effect of CPAP on Epworth Sleepiness Scale

Source First Author, Year	Fixed Effect Weight	Random Effect Weight	Study Estimate	Lower 95% CL	Upper 95% CL
Engleman 1997	0.96	0.67	-0.10	-2.10	1.90
Redline 1998	1.71	0.97	1.10	-0.40	2.60
Engleman 1999	1.96	1.05	3.00	1.60	4.40
Monasterio 2001	1.96	1.05	1.2	-0.20	2.60
Barnes 2002	5.17	1.57	0.40	-0.46	1.26
Barnes 2004	8.16	1.77	1.00	0.31	1.69
Marshall 2005	1.21	0.79	2.40	0.62	4.18
TOTAL			**1.2**	**0.5**	**1.9**

Figure 3.2: Meta-Analysis of the Effects of CPAP on the Epworth Sleepiness Scale



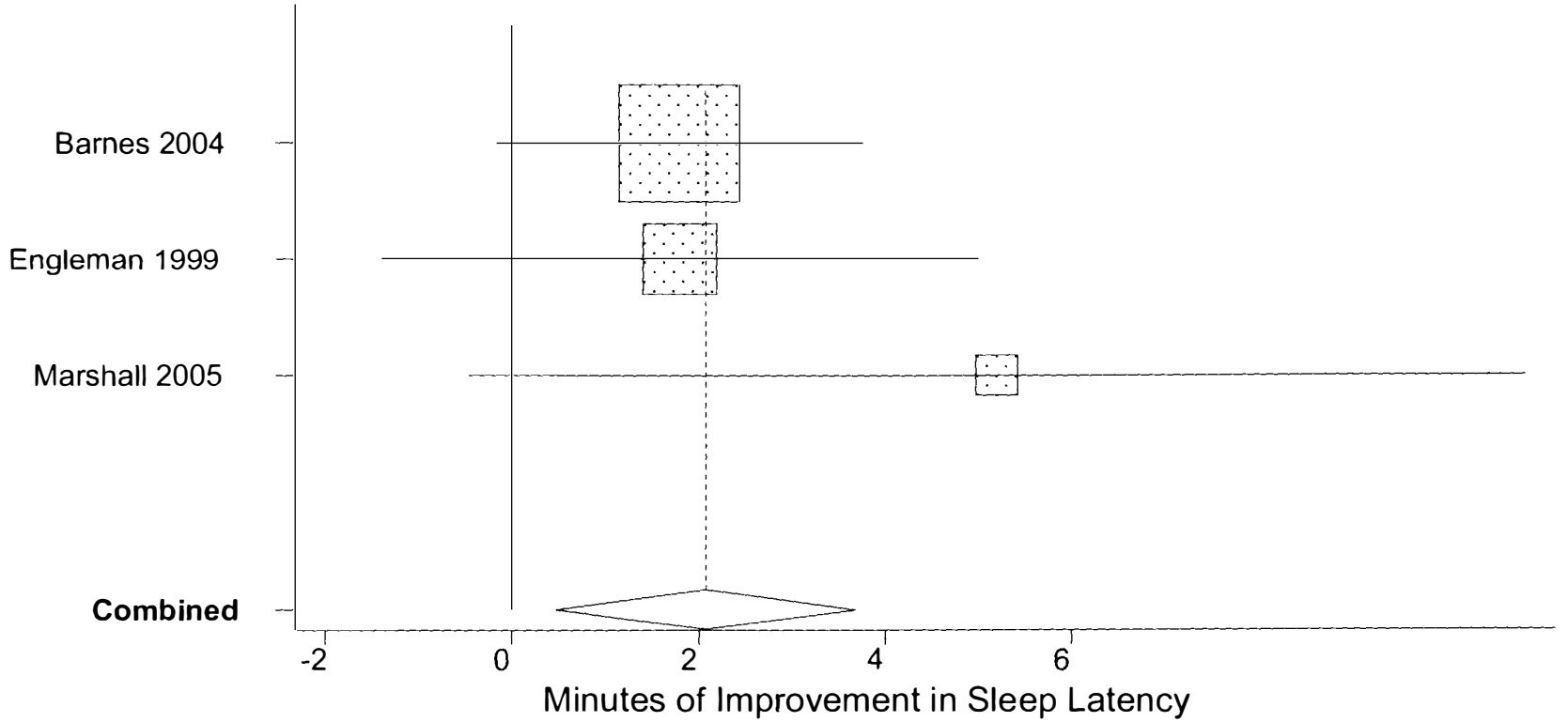
Maintenance of Wakefulness Test

The Maintenance of Wakefulness Test (Figure 3.3) was used in three studies and was found to improve by 2.1 minutes due to CPAP therapy (95% CI= 0.5 – 3.7, p=0.011), using both fixed and random methods. This agreement is because MWT effects were not heterogeneous (Q=1.3, df=2, p=0.53). Only two of the studies measured MWT at baseline, and were thus able to quantify the placebo effect. Barnes et al.[90] found a significant 2.7 minute worsening in MWT on placebo. Marshall et al.[38] found a non-significant 3.1 minute worsening whilst on placebo. These worsening sleep latencies contributed to the net effect of CPAP shown in Figure 3.3

Table 3.5: Effect of CPAP on MWT Latencies

Source First Author, Year	Fixed Effect Weight (Mins)	Random Effect Weight (Mins)	Study Estimate (Mins)	Lower 95% CL (Mins)	Upper 95% CL (Mins)
Engleman 1999	0.38	0.38	1.80	-1.40	5.00
Barnes 2004	1.00	1.00	1.80	-0.16	3.76
Marshall 2005	0.12	0.12	5.20	-0.46	10.86
TOTAL			**2.1**	**0.5**	**3.7**

Figure 3.3: Meta-Analysis of the Effects of CPAP on the Maintenance of Wakefulness Test



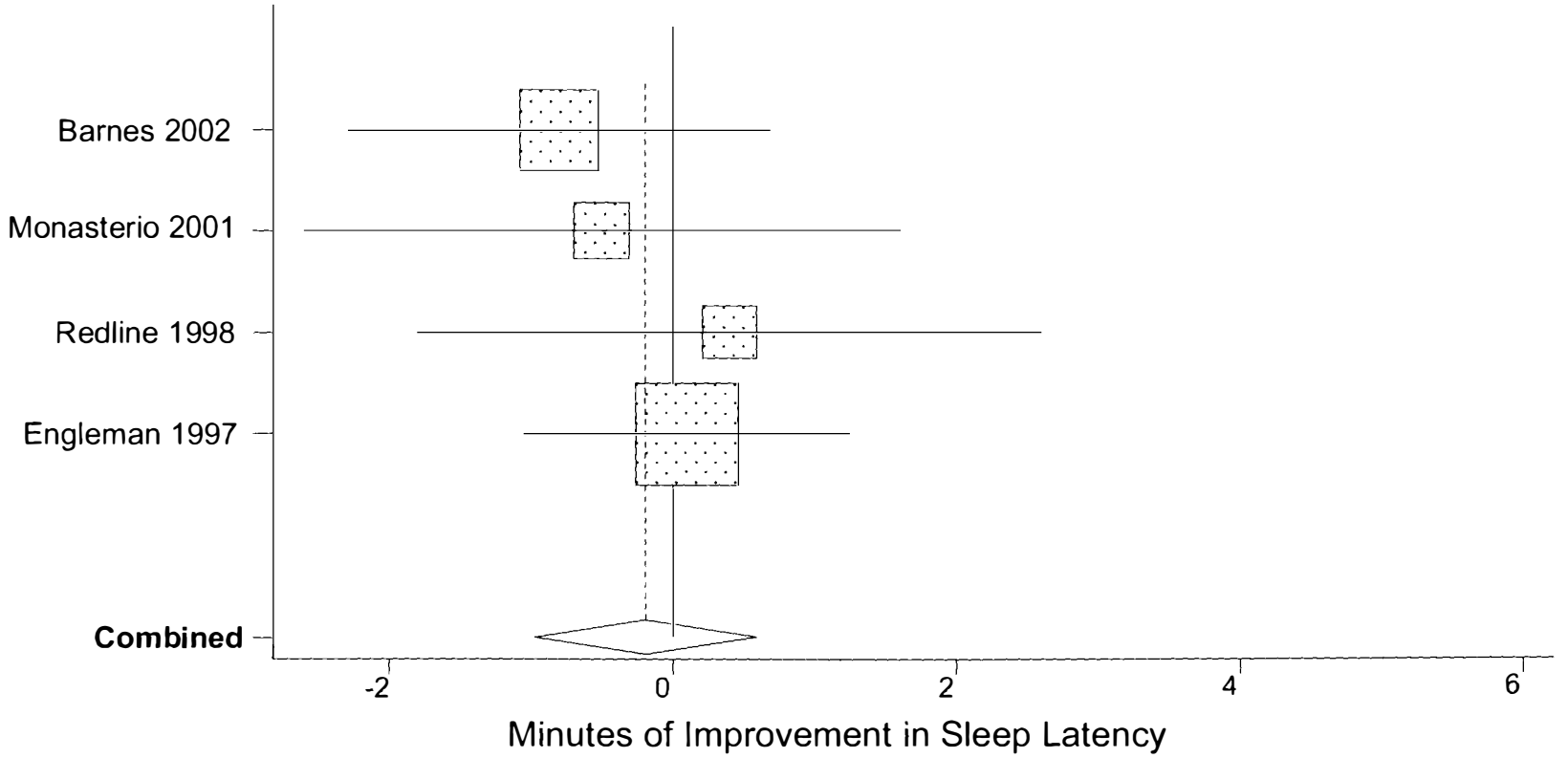
Multiple Sleep Latency Test

The MSLT (Figure 3.4) was used in four studies. Meta-analysis revealed a non-significant worsening of 0.2 minutes due to CPAP therapy, after controlling for placebo effects, using both fixed and random methods (95% CI -1.0 – 0.6 mins, p=0.74). Again, this agreement is because MSLT effects were not heterogeneous (Q=1.2, df=3, p=0.74). When the effects of CPAP on MSLT and MWT latencies were compared (-0.2 vs. 2.1 minutes), there was a significant difference (Meta Regression, Z=2.5, p=0.013).

Table 3.6: Effect of CPAP on MSLT Latencies

Source First Author, Year	Fixed Effect Weight (Mins)	Random Effect Weight (Mins)	Study Estimate (Mins)	Lower 95% CL (Mins)	Upper 95% CL (Mins)
Engleman 1997	2.90	2.90	0.10	-1.05	1.25
Redline 1998	0.79	0.79	0.40	-1.80	2.60
Monasterio 2001	0.87	0.87	-0.50	-2.60	1.60
Barnes 2002	1.73	1.73	-0.80	-2.29	0.69
TOTAL			**-0.2**	**-1.0**	**0.6**

Figure 3.4: Meta-Analysis of the Effects of CPAP on the Multiple Sleep Latency Test



Sensitivity Analyses and Publication Bias

Single trials were systematically removed one-by-one from each of the three analyses, to ascertain whether they had a significant influence on the total effect. No estimate of effect was overly reliant on the findings from one study (see Figure 3.5, Figure 3.6 and Figure 3.7). However, in the case of the MWT analysis (see Figure 3.6), when either the trail by Barnes et al. or Engleman et al. (1999)[90, 184] were removed, the 95% confidence intervals began to straddle zero due to reduced power. No evidence of bias was observed using funnel plots or either Begg's or Egger's tests for publication bias, for any of the three sleepiness measures (all $p>0.1$, see Figure 3.9 and Figure 3.10). However these tests for bias need to be interpreted with caution, due to the combination of only a small number of trials.

Figure 3.5: Effects of Study Omission on Epworth Estimate

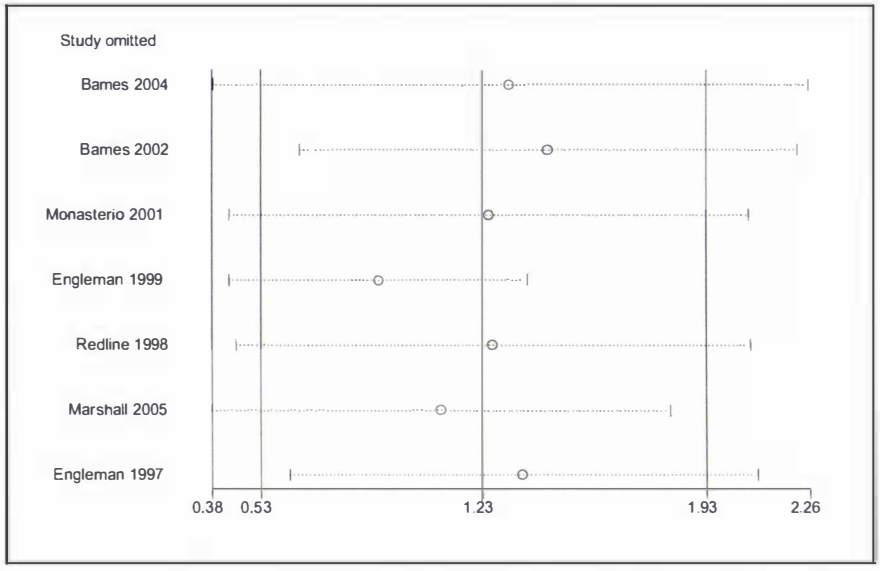


Figure 3.6: Effects of Study Omission on MWT Estimate

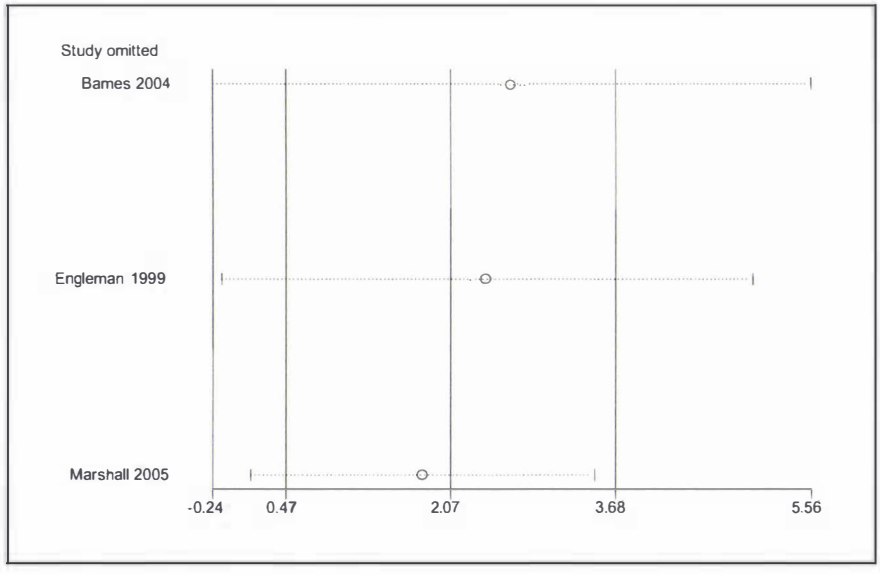


Figure 3.7: Effects of Study Omission on MSLT Estimate

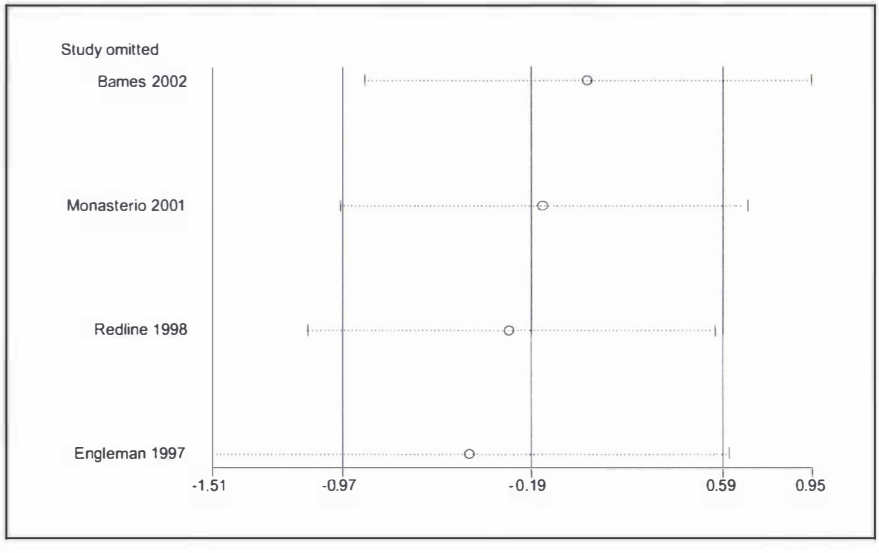


Figure 3.8: Begg's Plot for Publication Bias- Epworth Scores

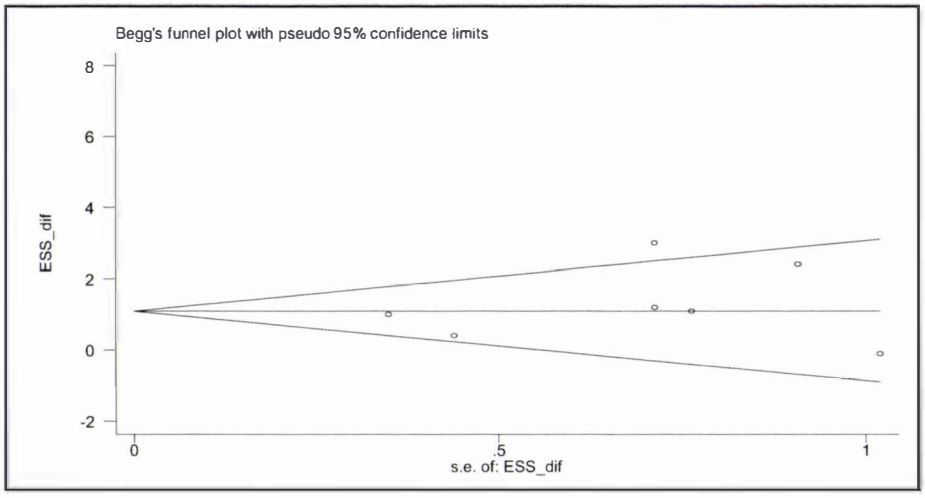


Figure 3.9: Begg's Plot for Publication Bias- MWT

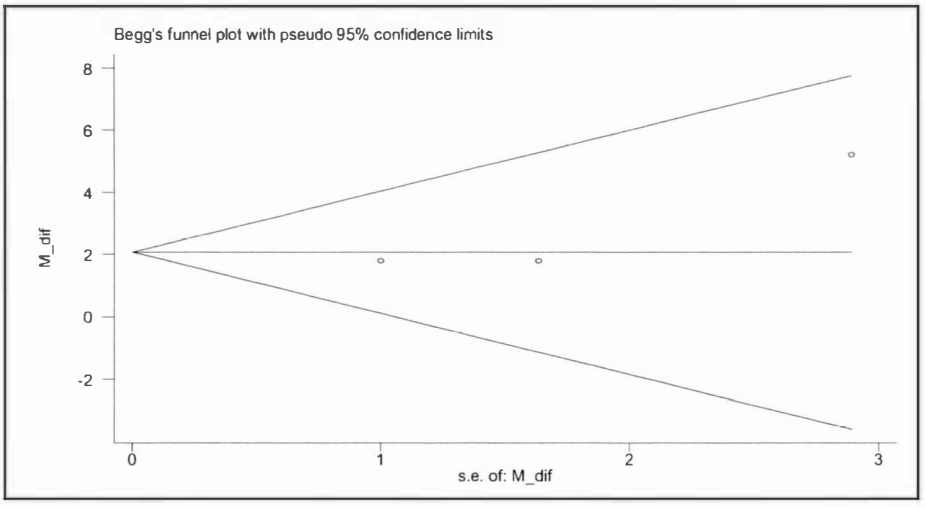
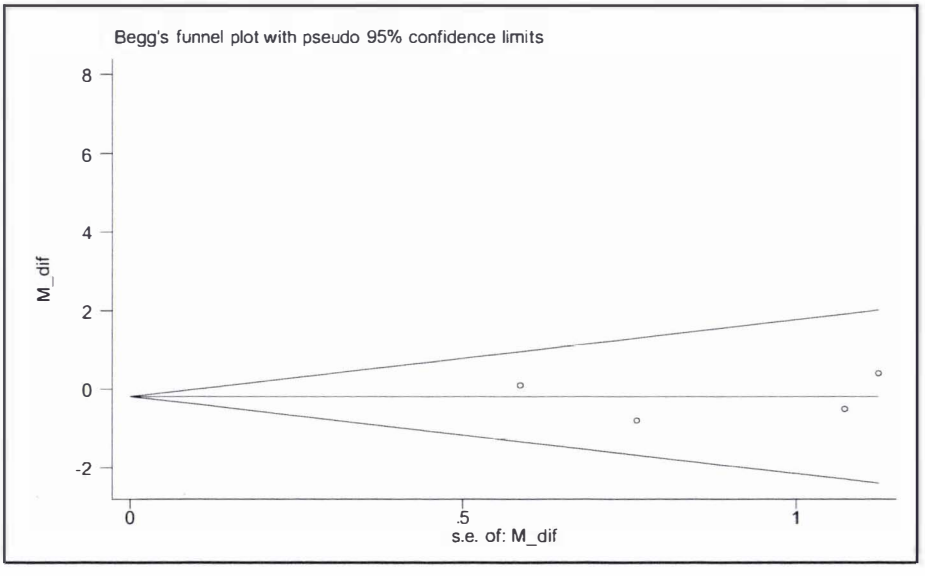


Figure 3.10: Begg's Plot for Publication Bias- MSLT



Size of the Effects

Epworth Sleepiness Scores improved by 1.2 points compared to a background standard deviation of 4.5 points[218], giving an effect size of 0.27. Maintenance of Wakefulness scores improving by 2.1 minutes compared to background standard deviation of 10.2 minutes, giving an effect size of 0.21. These effects are small (0.20-0.50) bordering on insignificant (<0.20)[112].

3.4 CONCLUSIONS

These analyses show, for the first time, that CPAP significantly reduces sleepiness, as measured by the Maintenance of Wakefulness Test and the Epworth Sleepiness Scale, in patients with mild-moderate OSAS, but that the effect is small. Sleepiness as measured by the Multiple Sleep Latency Test is not improved in these patients. In Chapter 5 the findings from this study will be presented in context and integrated with the current evidence base.

Matters to be discussed are:

- The small size of the effects and their potential clinical irrelevance.
- Irregularities in the effects of CPAP on Maintenance of Wakefulness Test latencies.
- Potential explanations of Epworth Score heterogeneity between studies.
- Relationship to, and differences from, the previous meta-analysis.
- The different sensitivity to treatment of the MSLT and MWT.
- The lack of cost-effectiveness data to support using treatment with marginal effectiveness.
- Large studies in progress that will be relevant to this study.

CHAPTER 4
PILOT BLINDED RANDOMISED CONTROLLED TRIAL OF NOVEL SELF-MODULATING CPAP (C-FLEX) COMPARED TO STANDARD CPAP.

4.1 INTRODUCTION

The acceptability of CPAP therapy can be low, with 5-50% of patients abandoning treatment within the first week, and only around 75-88% of those remaining able to maintain treatment over 3 years[Engleman, 2003 #412]. Between non-acceptance, and those patients who tolerate the device for their entire sleep period, are a significant proportion of patients with compliance levels less than their total sleep time[Kribbs, 1993 #286]. These patients may continue to have excess risk of cardiovascular disease, cognitive dysfunction, and excessive daytime sleepiness, because a proportion of their sleep disordered breathing is not controlled.

Patients overestimate their true compliance by more than an hour a night when self-reports are compared to objective measurement[Kribbs, 1993 #286]. Thus, objective compliance monitors, which can determine whether the mask is being worn or not, are now the standard method for measuring compliance with CPAP. Studies based on objective compliance have confirmed a wide range of underutilisation of CPAP. Underutilisation has been shown to be associated with a range of non-modifiable risk factors including lower AHI, lower baseline sleepiness level, lower age, and lower BMI[Engleman, 2003 #412]. Behavioural interventions, particularly intensive patient support and follow-up have been found to moderately aid compliance[Engleman, 2003 #412]. There are also a number of technological advancements, or additions to CPAP, that are now being marketed by manufacturing firms as compliance enhancers. These

add-ons, or adjustments to PAP delivery, present a clear potential for technological drift. Marginal, ever more expensive add-ons can, and have, become entrenched in standard clinical practice without rigorous testing in RCTs.

Technological Device Modifications Aimed At Improving Compliance

Humidification of Positive Airway Pressure

OSAS patients commonly complain of airway dryness and associated symptoms when being established on CPAP (or other PAP therapy)[Engleman, 2003 #412]. This airway dryness is, with some justification, often blamed on CPAP therapy and may explain some of the poor compliance with CPAP. CPAP devices now commonly come with built in, or add-on heated humidifiers, that attempt to alleviate upper airway dryness.

Two randomised controlled trials have assessed the effects of heated humidification on compliance. Massie and colleagues[Massie, 1999 #615], have shown that compliance during a three-week trial could be increased when heated humidity (5.52±2.1 hrs/night) was used, compared to no humidity (4.93±2.2 hrs/night, difference 35 mins, $p<0.01$). Engleman and Wild calculated the effect size of this intervention to have been small, at 0.3[Engleman, 2003 #412]. Heated humidification was not significantly better than cold pass-over humidity (5.15±1.9 hrs/night). However, the study was underpowered to detect the observed difference between heated and cold pass-over methods. Heated humidification is usually favoured because it has been shown to humidify air more effectively[WIEST, 2000 #876].

Neill and colleagues[Neill, 2003 #660] reported similar results when they compared heated humidified CPAP compliance with the compliance on CPAP with placebo humidification. They reported an increase of 0.4 hrs/night (24 mins) on heated humidified CPAP above the 5.3 hrs/night (SD 1.7) on placebo humidified CPAP (p for diff=0.03). This difference, when compared to a background standard deviation of 1.6 hours/night, gives a small effect size of 0.25[Kazis, 1989 #344]. In neither study were the increased levels of compliance associated with greater reductions in daytime sleepiness.

CPAP Pressure Ramping

Many CPAP devices now come with pressure ramps as a standard feature. These allow the patient to start the CPAP device at a lower pressure which then gradually reaches full therapeutic pressure in 5-30 minutes, presumably after the patient has fallen asleep. These device add-ons have barely been studied systematically and have not yet been shown to increase patient comfort. They are clearly a leading contender for an example of technological drift. Grunstein, for instance, wonders whether, in some instances, such add-ons are "merely cosmetic marketing ploys"[Grunstein, 2005 #3].

Automatically Titrating Positive Airway Pressure Devices (AutoPAPs)

High pressure from PAP devices is often thought to limit compliance, but there is little empirical evidence of this[Berry, 2002 #320]. CPAP pressure is constant, but the pressure requirements to hold open a collapsible airway differ according to sleep stage and sleeping position, so patients are often receiving more pressure than they require. This is often the case around sleep onset and offset where the patient is, at least partially, conscious, and able to pull the mask off.

AutoPAP devices were developed to titrate pressure requirements, breath-by-breath, to avoid application of more positive pressure than was strictly required to splint open the patient's airway. Without much supporting evidence, these devices are believed to increase compliance compared to standard CPAP, and are widely used because of this purported benefit[Berry, 2002 #320].

Only after the widespread adoption of autoPAPs did clinical trials begin to appear in the literature. The trials indicated that for most patients, CPAP was of similar effectiveness. The lack of superiority in clinical trials and the 2-3 fold increase in the cost of an autoPAP device over a CPAP device, present a good example of a technological drift.

A recent bench-top study aimed to simulate various types of sleep disordered breathing, and tested the responses of 5 different autoPAPs[Farre, 2002 #323]. There was considerable variability in the abilities of the devices to respond to a range of simulated sleep disordered breathing, and none of the devices was wholly satisfactory. An autoPAP device that is effective for one patient may not be effective for another, and some patient groups might be particularly badly affected by sudden pressure increases due to inadequate algorithms. Furthermore, clinical trials of one particular device do not necessarily have relevance for the evaluation of a different autoPAP device.

In 2002, the American Academy of Sleep Medicine (AASM) reviewed the published studies of the various autoPAPs[Berry, 2002 #320] and issued guidelines as to their use[Littner, 2002 #321]. The AASM found that published accounts of autoPAP performance indicated that it was generally equivalent to CPAP in terms of controlling

AHI and reducing daytime sleepiness. Conflicting evidence was found regarding the hypothesis that autoPAPs increased compliance compared to CPAP. Six studies had found between 1.4 and 0 hours of additional compliance with autoPAP (3 found compliance benefit, 3 found no benefit over CPAP). Thus, the reviewers were unable to conclude that autoPAPs, as a class, increased compliance[Berry, 2002 #320]. Nor were they able to conclude that the reduction in mean or median pressure was a likely cause of increased compliance.

More recently, the effects of autoPAPs, compared to CPAP, have been combined into a meta-analysis[Ayas, 2003 #620]. Nine studies were identified. There were no advantages to the control of SDB compared to CPAP. Nor was there any advantage to reducing sleepiness (Epworth). No effect on improving compliance was found, despite the reduction in mean pressure across the night of 2.2cmH2O.

One subsequent crossover trial, has investigated the use of autoPAP in patients with high variability in their pressure requirements. Despite similar control of the AHI and similar compliance levels, the Epworth Sleepiness scale dropped by 1 point more on autoPAP than on CPAP[Noseda, 2004 #741].

BiLevel Positive Airway Pressure

BiLevel PAP is differentiated from standard CPAP by the application of two pressures: one pressure at inhalation, (IPAP) and a lower one at exhalation (EPAP, see Figure 1). BiLevel devices are designed to sense the beginning of inspiration, switching between IPAP and EPAP, and are used mainly to treat patients in type II respiratory failure. BiPAPs are also routinely used to treat severely obese OSAS patients who hypoventilate or those with overlapping OSAS and chronic obstructive pulmonary

disease (COPD)[Resta, 1998 #431;Schafer, 1998 #432;Grunstein, 2005 #3]. However, it seems no studies have systematically investigated the benefits to patients with restricted lung disease[Grunstein, 2005 #3]. Given that these devices usually cost around 10 times as much as standard CPAP, BiLevel PAP is less often chosen for treating uncomplicated OSAS. In the only RCT to compare CPAP and BiPAP compliance there was no difference after 12 months of therapy[Reeves-Hoche, 1995 #908].

A device that was a prototype of the BiFlex device (BiPAP but with scalloped inhalation and exhalation pressure profile rather than a stepped profile) has been tested in a randomised trial against CPAP[Gay, 2003 #629]. As a first-line therapy for uncomplicated severe OSAS, but again it offered no compliance benefits over standard CPAP.

CPAP That Modulates Pressure During Exhalation (C-Flex)

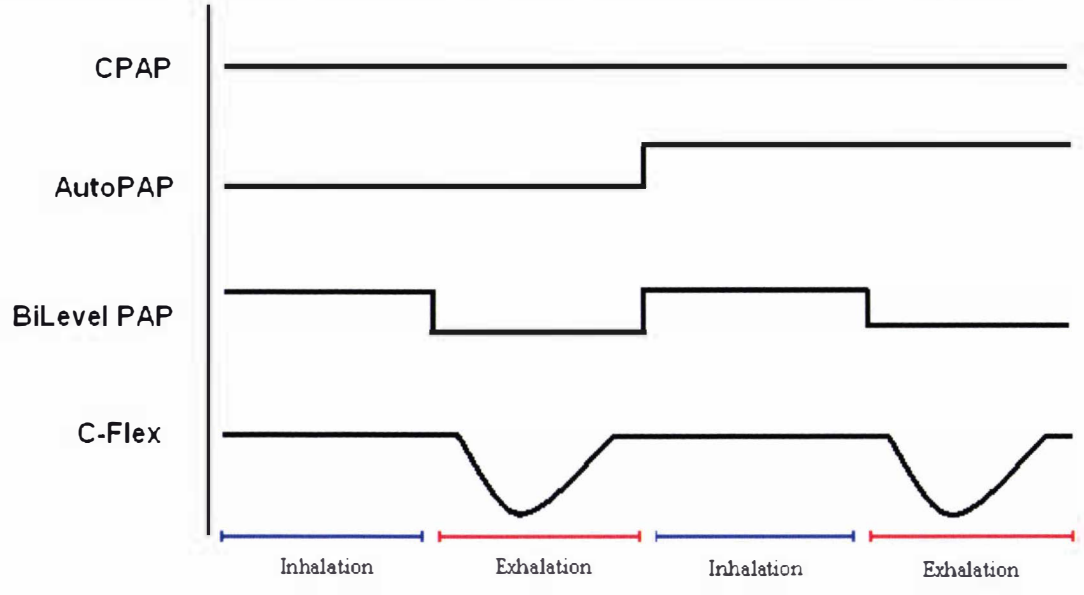
Currently only one manufacturer produces a CPAP device that modulates pressure within each breath, the REMstar CPAP devices with C-Flex (Respironics Inc, NC, USA). C-Flex, reduces pressure in response to exhalation. In contrast, BiPAP uses exhalation as a trigger to step abruptly to a lower pressure. As a result, the pressure applied by the device, over time, looks scalloped. (see Figure 4.1 for comparison to other modalities).

The manufacturers of C-Flex have web-published 3 small, apparently non-peer reviewed, CPAP equivalence trials (Appendix 2 and www.respironicsremstar.com). These rely on patient testimonials and poorly-designed polysomnographic protocols to support the use of C-Flex. One non-randomised study compared the compliance of a group of 5 patients on CPAP and 4 patients on C-Flex. Those on C-Flex used the

machine 5.83 hours/night compared to 4.74 hours/night for those on CPAP. However, the CPAP group had been using the device for a mean 137 days, while the C-Flex group had a mean 64 days of use. The differing compliance rates could be due to declining compliance over time, or some other non-randomised factor. No statistical analyses are presented. Thus there is a need for an independently conducted, blinded randomised controlled trial that compares compliance on the C-Flex and CPAP modes on the REMStar device.

A non-randomised study appeared in Chest in June 2005, after the study being described here had already been completed[Aloia, 2005 #905]. It indicated that C-Flex resulted in a substantial increase in compliance over that achieved with CPAP. The implications of this study are discussed in Chapter 5.

Figure 4.1: Pattern of Pressure Applied by Various Positive Pressure Devices



Pressure deviations are not necessarily to scale and the magnitude of change is not necessarily comparable between PAP types. Pressure traces indicate the direction and immediacy of pressure shift only. AutoPAPs can also change pressure down. AutoPAPs will also tend to change pressure in response to upper airway resistance/blockage rather than strictly at the start of inhalation.

4.2 METHODS

Study Design, Randomisation, and Blinding

This study was a parallel arm, single blinded, randomised controlled trial comparing the effects of C-Flex and CPAP, on compliance and daytime sleepiness after 4 weeks of treatment. Randomisation to treatment with CPAP or C-Flex occurred after the first day of testing, but before pressure titration, or a combination diagnostic and titration study (split night study[Sanders, 2000 #870;Rodway, 2003 #749;McArdle, 2000 #845]). The sequence of allocation to treatment was determined by sleep technicians randomly picking one of a set of pre-prepared opaque envelopes containing the treatment allocation.

To maintain sequence unpredictability and balance group sizes, prospective Urn Randomisation using a (2,1) schedule was used[Schultz, 2002 #487]. An urn with 2 differently coloured paper clips was prepared. One clip was blindly withdrawn, and its corresponding treatment noted on a folded piece of paper in an opaque envelope which was then sealed. The clip was then replaced in the urn along with another coloured clip, which represented the other treatment. This was repeated until 20 envelopes had been produced. Because a number of patients who did not meet the entry requirements, were randomised and then withdrawn from the study, a further 5 envelopes were subsequently prepared using the same method. This method leads toward roughly equal group sizes whilst maintaining unpredictability of sequence and sample sizes.

Patients were blinded to treatment allocation. They were told that we were testing a new machine that had two different modes, to see which was the most comfortable, and that they would be receiving one of these modes. Patients were further reassured that one of these modes was not a placebo, and that we currently had no reason to think that either mode was an inferior treatment. They were also told that both treatments would abolish their SDB when the mask was worn correctly.

After two weeks of therapy, a compliance and mask fitting follow-up was conducted according to normal clinical practice, by a sleep technician (Sue Garrett) who was not blinded to treatment allocation. Mask fitting and other comfort problems could be dealt with at these meetings. Both treatment groups received the same levels of support/follow-up.

I undertook the collection of secondary outcome measures before and after 4 weeks, and was blinded to treatment allocation. Thus the trial was only single-blinded in terms of the primary outcome measure, somewhat double-blinded to the secondary outcome measures whilst they were being collected, and was additionally blinded during data analyses.

Pre Trial Power Calculation

The primary aim of the study was to find whether, prima facie, the one hour difference in compliance between treatments claimed by the manufacturers could be demonstrated, and to estimate variability in compliance to enable power calculations for further trials. The inter-patient variability in C-Flex compliance is unknown. Twenty

patients were thought adequate to make this estimate for future trial power calculations. The standard deviation for CPAP use is usually about 2-3 hours, but this measure of variability might not be appropriate given that compliance distributions are often bimodal (see Chapter 2 for an example). With ten patients in each group, the present trial was adequately powered to detect differences of about 1.3 standard deviations or more, and was thus adequately powered to detect only very large differences. If the standard deviation for CPAP compliance were 2 hours per night, then the present trial would be able to detect difference in compliance of at least 2.6 hours, 80% of the time.

Recruitment and Entry Requirements

English speaking adult patients (18+ years), referred from local sleep clinics and naïve to CPAP, were prospectively screened for study inclusion. Patients were enrolled to the trial by Dr Angela Campbell, Karyn O'Keeffe (Sleep Technician), or Marion Vickerman (WellSleep Receptionist and Patient Scheduler) if they had a preliminary diagnosis of severe OSAS from a Sleep Physician (Drs David Jones, Justin Travers or Alister Neill) or had already had severe OSAS (AHI 30+) confirmed by overnight polysomnography. All participants gave written informed consent to be in the trial, after I further introduced the study.

Patients were excluded if they;

- had a history of extreme somnolence requiring immediate treatment
- were shift-workers and/or were chronically sleep restricted (average Total Sleep Time $\leq$ 6 hr/night)
- took sedative, antidepressant, psychotropic or stimulant medications;
- consumed more than > 3 standard units/24hr of alcohol

- had caffeine dependency (felt unable to forgo caffeine on study days)
- had upper airway surgery since the diagnostic sleep study (if any such study)
- had clinically significant co-existing disease and/or additional sleep disorders.

Significant exclusionary co-morbidities included:

- heart failure
- severe diabetes (type 1 or 2)
- respiratory failure
- unstable angina
- organ failure
- any significant psychiatric disorders.

Significant exclusionary sleep disorders included:

- central sleep apnoea (and/or Cheyne-Stokes respiration pattern),
- narcolepsy,
- periodic limb movement disorder (>10/hour),
- obesity related hypoventilation,
- other disorders of REM.

Patients were later excluded if diagnostic polysomnography indicated an AHI less than 30/hr, or between 20 and 30/hr, but without an Epworth Sleepiness Score of at least 13, indicating severe daytime sleepiness.

Diagnostic Polysomnography

Polysomnographic measures and the measures of SDB were the same as described in Chapter 2. Studies were scored in 30 second epochs (Dr Angela Campbell, Michi Imazu, Karyn O'Keeffe or Deidre Sheppard) using accepted international criteria[Rechtschaffen, 1968 #318;Association, 1992 #334].

Continuous Positive Airway Pressure and C-Flex Systems

Therapeutic CPAP pressure was determined by manual titration by a WellSleep Technician (Karyn O'Keeffe, Deidre Sheppard, Helen Morgan or Michi Imazu) to abolish apnoeas and hypopnoeas, and reduce respiratory related arousals, and was later confirmed by a sleep physician (Drs Alister Neill or David Jones). The Respironics REMStar with C-Flex device was able to be set to either CPAP or C-Flex (2) mode, via a switch on the machine. The switch was sealed with an anti-tamper cover (designed and installed by Karyn O'Keefe and Dr Angela Campbell) to avoid accidental or deliberate changes in the mode of pressure delivery, by patients or staff. C-Flex mode can be supplied in three depths (1-3) of pressure dipping. The middle level (2) was used exclusively during this trial. Both treatment modes also came with heated humidification as standard.

Split Night Studies- Diagnosis and Titration on the Same Night

Split night studies are typically used where a strong clinical likelihood of severe OSAS is identified by the referring physician. The split night study begins with a

diagnostic phase to confirm the presence of severe OSAS. At Wellsleep, this usually lasts around 3 hours, and requires a full sleep cycle in both the supine and lateral positions to quantify the severity of SDB across all sleep stages and positions. CPAP titration is then initiated, if, by 3am, severe SDB has been confirmed. This allows sufficient time for an accurate determination of the pressure required to splint the airway through the relevant sleep stages and positions.

Split night studies are widely used around the world (including Australasia) as they are less expensive and minimise the time between diagnosis and treatment for patients with severe OSAS[Elshaug, 2005 #878]. Split night studies do not appear to negatively affect patients with severe OSAS, as titration accuracy and longer-term compliance with treatment seem comparable to full night titration[McArdle, 2000 #845;Yamashiro, 1995 #871;Rodway, 2003 #749;Sanders, 2000 #870]. However, the evidence supporting the validity of split-night studies is not universally accepted as adequate[Grunstein, 2005 #3].

Daytime Procedure

Patients were instructed to arrive at 1230 hours on both days of testing, in order to have completed orientation and electrode application by the start of daytime testing at 1330. The testing timetable can be found in Table 4.1.

Table 4.1 Daytime Testing Timetable

Time of Day	Outcome Variable
1230	Patient Orientation.
	Placement of scalp and face electrodes
1330	Psychomotor Vigilance Task 1
1400	Maintenance Of Wakefulness Test 1
1450	New Zealand National Sleep Survey
1530	Psychomotor Vigilance Task 2
1600	Maintenance of Wakefulness Test 2
1700	Patient Leaves (if visit 2).
	Blinded daytime data collector leaves (NSM)
Evening	Overnight CPAP or C-Flex titration (if visit 1)

Outcome Measures

Compliance

Compliance was objectively measured by an internal pressure sensor in the device and expressed as the average number of hours use per night averaged over the 4-week treatment period (Respironic REMStar Pro with C-Flex, Encore Pro Smartcard, Respironics, NC, USA). The compliance measure was the amount of time that the mask was being worn correctly and the machine had sensed effective pressure delivery.

A limited number of secondary outcome measurements were collected and were chosen for comparability to other RCT's.

Epworth Sleepiness Scale

In this study the Epworth Sleepiness Scale[Johns, 1991 #54] was embedded in the questionnaire used by Harris[Harris, 2003 #889], and was given to patients to

complete on both study days. Harris' questionnaire was employed to disguise the Epworth scale, which the patients may have subjected to twice earlier during clinical referral for polysomnography.

Modified Maintenance of Wakefulness Test

A modified Maintenance of Wakefulness Test (modMWT) was employed to objectively measure patients' ability to remain awake in a soporific environment(see Appendix 1 and Chapter 2 for review)[Mitler, 1982 #72;Doghramji, 1997 #71;Banks, 2004 #558]. The test is normally repeated 4 times throughout the day (typically 1000, 1200, 1400, 1600, and possibly 1800), however only two tests were employed here, at 1400 and 1600.

Psychomotor Vigilance Task

The Psychomotor Vigilance Task (PVT)[Dinges, 1985 #295] is a handheld device that tests simple primary reaction times (see Appendix 1 and Chapter 2). Two ten-minute tests were administered at 1330 and 1530 each day, with a 1-minute practice session before each test. Mean reaction times and lapses (reactions longer than 500ms) were measured.

Data Handling and Statistical Analyses

All data were double entered into Excel (Microsoft Corp, WN, USA) by Nathaniel Marshall, except the compliance data and the treatment allocation which, to preserve blinding, were entered by the staff at WellSleep. Data analyses were conducted in SPSS (v.12.0.1, SPSS Inc. Chicago, USA). Modified MWT and PVT data were

averaged within each day. Each individual had two data points, one at the beginning and one at the end of the trial, which also yields a net treatment effect. The effects of treatment, within and between arms, were analysed. All statistical tests were regarded as significant when $p<0.05$. Continuous data were tested for normality using the Kolmogorov-Smirnov test and visual checking of histograms of the data under normal curves. Equality of variance between treatments groups (for t-tests) was tested with Levene's test.

Within-arm treatment effects were tested using paired t tests (before and after pairings) for the following variables:

- PVT Mean Reaction Time
- PVT Lapses (RT>500ms)
- Epworth Sleepiness Scale
- Maintenance of Wakefulness Test
- All polysomnographic measures

Differences between CPAP and C-Flex treatment effects were tested using independent t-tests for the following variables:

- Compliance
- PVT Mean Reaction Time
- PVT Lapses (RT>500ms)
- Epworth Sleepiness Scale
- Maintenance of Wakefulness Test
- All polysomnographic measures

Tests for differences where normality was questionable were repeated with the Mann-Whitney U for confirmation:

- PVT Mean Reaction Time
- PVT Lapses (RT>500ms)

Normally-distributed continuous variables were correlated with each other using Pearson's r statistic.

Effect sizes were relied upon in this pilot trial, because statistical significance was unlikely to be shown. Effect sizes were used to quantify the magnitude of treatment-induced changes, and were calculated by dividing the mean difference by the standard deviation of that measurement at baseline. Small effect sizes are between 0.20 and 0.50, medium 0.50 and 0.80, and large effect sizes >0.80 standard deviations[Kazis, 1989 #344].

4.3 RESULTS

Patient Characteristics, Randomisation and Retention

Patient characteristics can be found in Table 2. Patients were typical for clinically-referred, publicly funded patients with suspected OSAS. They were predominantly male, middle-aged, overweight, and sleepy. The self-identified ethnicities of the participants were as follows, 2 Māori, 4 Samoans, 10 New Zealand European/ Pakeha, 1 Cook Islander, and 2 others (1 Filipino, 1 Bangladeshi). Participants also reported being treated for the following co-morbidities: asthma (3/18),

hypertension (7/18), heart trouble (1/17), diabetes (2/18), stroke (1/17), and hyper- or hypo-thyroidism (3/18). Eight out of eighteen, reported being eligible for a community services card. In New Zealand this card allows low-income earners access to cheaper healthcare and other public services, and it can be used as a crude individual marker of socioeconomic deprivation.

Five patients were randomised to treatment before split-night studies, but were later withdrawn because they did not meet the polysomnographic entry requirements. Four had AHIs that were too low for study inclusion (one of these also had Periodic Limb Movement Disorder). The remaining patient had a high AHI, but received supplementary oxygen during the night for obesity hypoventilation syndrome, and was excluded because of this.

Two patients who initially enrolled in the study failed to present at the daytime testing. They did attend a split-night study and would have been eligible for this study, but were not randomised to treatment because they had not given informed consent. These two patients did not seem to be different from those who were randomised to treatment (see Table 2 for comparison with study sample).

Seventeen of the 19 patients included in analyses received split-night studies. One had an in-lab diagnostic study and then an in-lab titration study. The other had an in-lab diagnostic study that was unsatisfactory, followed by a full in-home diagnostic polysomnography, and then an in-lab titration, once severe OSAS was confirmed.

Table 4.2: Patient Characteristics

Patient Characteristic	Participants Mean (SD) n=19	Non-participants n=2
Number of Females	4 (21%)	0 (0%)
Body Mass Index kg/m^2	38.8 (9.5)	39.7 and 45.1
Age (Years)	46 (range 20-63)	42 and 44
Epworth Score (/24)	14.3 (4.1)	14 and 19
Sleep Efficiency (%)	67 (21)	85.3 and 95.1
Arousal Index (/hr)	52 (26)	42.5 and missing
AHI	78 (33)	49.7 and 87.4
CPAP Pressure (cmH_2O)	13.7 (2.9)	Both 9
ModMWT (mins)	18.5 (9)	Unknown
Mean Reaction Time (ms)	498 (366)	Unknown
Lapses (RT>500ms)	14.9 (19.8)	Unknown

SD=Standard Deviation. AHI- Apnoea Hypopnoea Index. ModMWT= Modified Maintenance of Wakefulness Test

Ten patients were randomised to CPAP treatment and 9 to C-Flex treatment. A comparison of these groups can be seen in Table 3. There were no post-randomisation dropouts (other than the 5 withdrawals due to ineligibility). Two patients failed to attend their second data collection day. One failed to turn up for the first day of treatment, but had already given his informed consent by post, after having the study explained to him by phone. He was randomised to treatment before his split-night study. These 3 patients are all included in the primary analyses, under the intention to treat principle, as their compliance levels were available and they had given informed consent to be in the trial. The patient who was unable to attend the first day completed

an Epworth Sleepiness Scale before his split-night study and thus could only be included in the Epworth analyses. The two remaining patients, who did not attend their second day of testing, were excluded from all the secondary data analyses. Thus analyses for the primary outcome included 19 patients, 17 patients are included in the Epworth Analysis and 16 each in the modMWT and PVT analyses.

Effects of CPAP and C-Flex on Sleep Disordered Breathing

CPAP and C-Flex had similar effects on indices of SDB during the titration phase. As expected, both treatments reduced the AHI and improved oxygen desaturation indices[Sullivan, 1981 #4] (see Table 4).

Table 4.3: Comparison of patients randomised to C-Flex with those randomised to CPAP

Patient Characteristic	CPAP Group Mean (SD) (n=10)	C-Flex Group Mean (SD) (n=9)
Gender (Female)	2/10 (20% Female)	2/9 (22% Female)
Body Mass Index kg/m^2	37.7 (9.6)	40.0 (9.8)
Age (Years)	48.7 (Range 29-64)	44.9 (Range 20-61)
Epworth Score (/24)	14.5 (4.6)	14.0 (3.8)
Ethnicity	2 Māori 1 Samoan 6 NZ European/Pakeha 1 Bangladeshi	3 Samoan 4 NZ European/Pakeha 1 Cook Islander 1 Filipino
Eligibility for Community	5 Yes	3 Yes
Services Card	5 Yes	
Self Reported Sleep Duration	6.6 (1.2)	6.8 (1.5)
Sleep Efficiency (%)	66.3 (19.9)	68.0 (23.1)
Arousal Index (/hr)	29.7 (31.6)	53.7 (29.0)
Apnoea Hypopnoea Index	77.5 (35.1)	77.7 (32.0)
Prescribed CPAP Pressure	13.3 (range 8-17)	14.1 (range 8-17)
MWT (Mins)	18.2 (8.15)	18.9 (10.5)
Mean Reaction Time (ms)	460 (334)	547 (420)
PVT Lapses (RT>500ms)	12.4 (15.9)	18.1 (24.7)

No hypothesis testing for differences were included in this table, as differences could only occur by chance with a randomisation process. MWT= Maintenance of Wakefulness Test. PVT= Psychomotor Vigilance Task. RT=Reaction Time.

Table 4.4: Effects of Treatment on Polysomnographic Markers of Sleep Disordered Breathing

Polysomnographic Measure	CPAP		C-Flex		p for difference between treatment effects
	Baseline Mean (95% CI)	During Mean (95% CI)	Baseline Mean (95% CI)	During Mean (95% CI)	
Sleep Efficiency (%)	66.3 (54-79)	74.3 (62-87)	68.0 (55-81)	77.7 (64-91)	0.79
Arousal Index (/hr)	29.7 (15-45)	30.1 (15-45)	53.7 (37-71)	14.3 (-2-30)**	0.20
AHI (/hr)	77.5 (61-94)	21 (4-38)***	77.7 (60-95)	11.7 (-5-29)***	0.59
Min O_2 Saturation -					
NREM	76.2 (69-83)	82.8 (76-90)	76.7 (70-84)	87.3 (80-94)*	0.27
REM	71.3 (64-78)	89.3 (83-96)***	75.2 (67-83)	88.6 (82-95)*	0.86

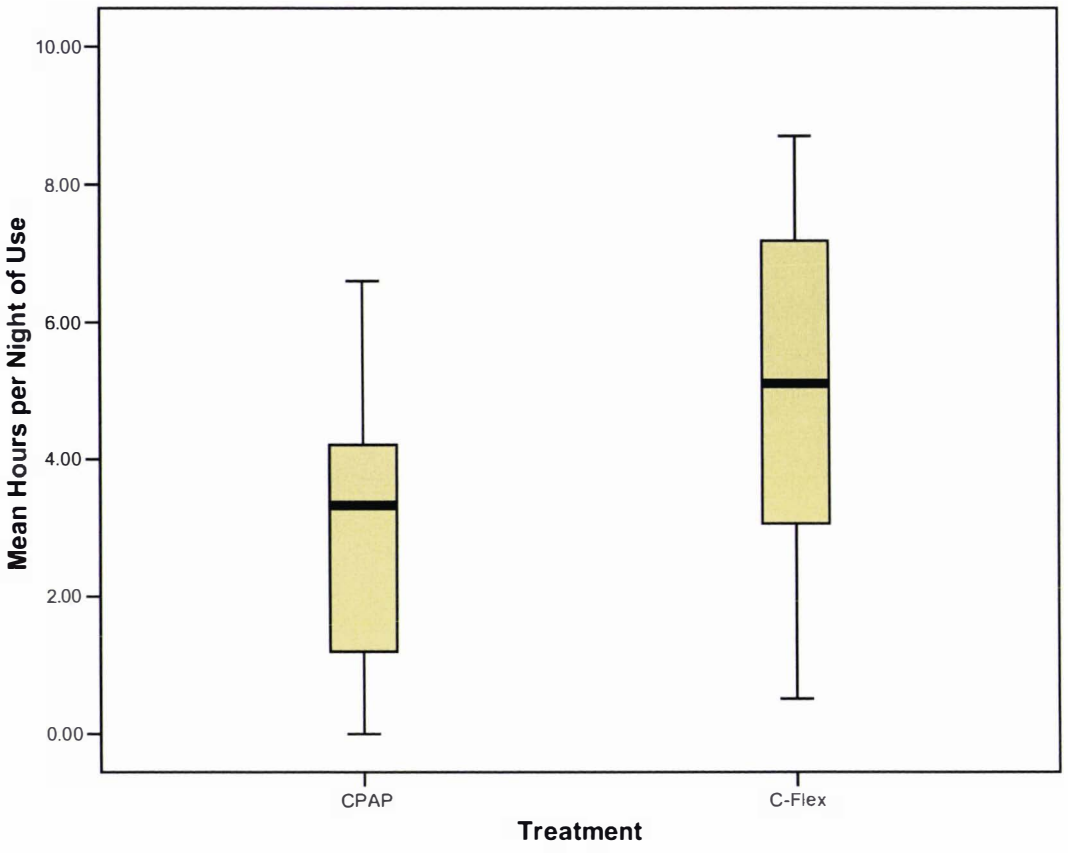
Differences between and within treatments were normal and differences were tested for with independent samples t-tests. *p<0.05, **p<0.01, ***p<0.001. Figures during treatment are from the whole period of titration, and thus contain large amounts of time where less than optimal pressure was being delivered. This is why control of SDB seems to be relatively poor.

Primary Outcome: Compliance with Therapy

The primary analysis compared the average nightly compliance on CPAP with the compliance on C-Flex, under an intention to treat principle. Both compliance distributions were approximately normal when tested with the Kolmogorov-Smirnov test (both $p>0.7$), and the variability was not different between treatments ($F=1.9$, $p=0.36$). CPAP compliance averaged 3.0 hours per night (SD=2.1 hours, range 0-6.6 hrs/night), whilst C-Flex compliance averaged 4.7 hours per night (SD=2.9 hours, range 0.5 to 8.7 hrs/night). This difference, whilst large, was not statistically significant ($t=1.52$, 95% CI -0.7 to 4.1 hrs/night, $p=0.15$, see Figure 4.2). Three of 10 patients (30%) used CPAP, and 6 of 9 patients (67%) used C-Flex more than 4 hours per night on average. The effect size of 0.68 is moderate. The 1.7 hours estimated additional compliance is much larger than the average 30 minute improvement seen with the addition of heated humidification to CPAP[Massie, 1999 #615;Neill, 2003 #660].

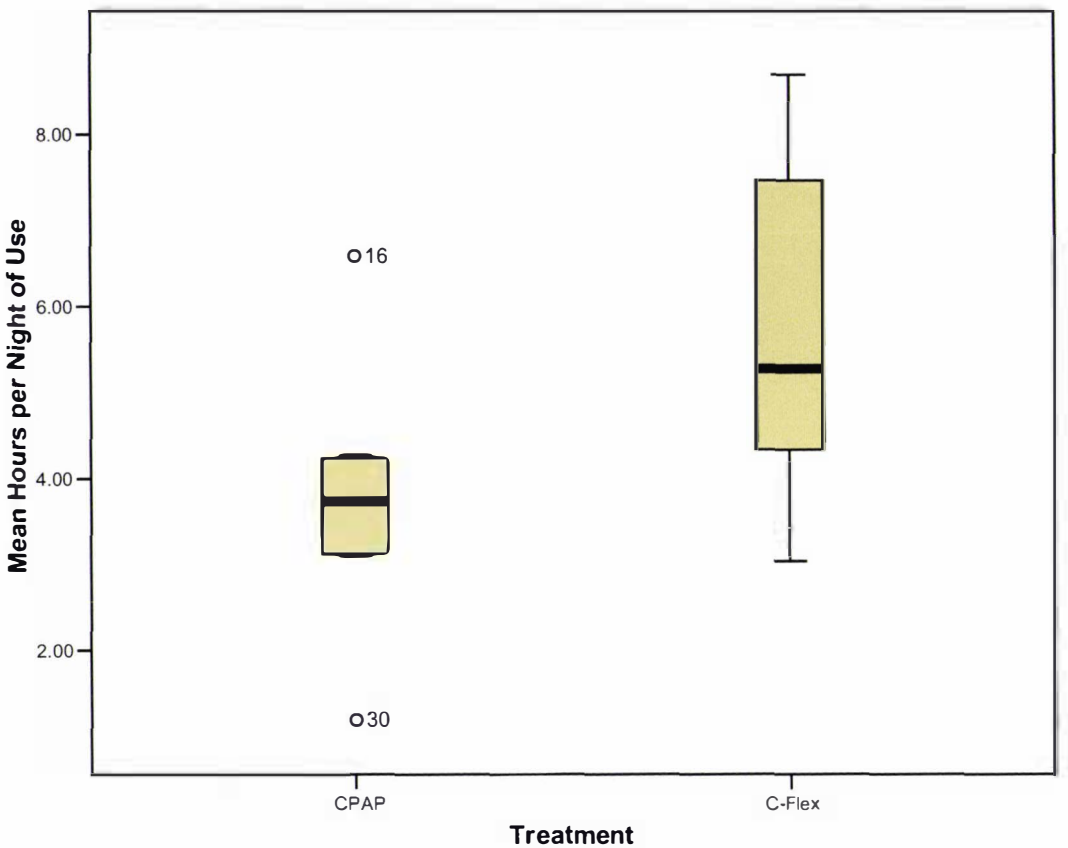
A post-hoc power calculation indicated that approximately 33 patients per group would be required to give an 80% likelihood of detecting a 1.7 hour difference as being significant (alpha=0.05) with the observed variability.
http://calculators.stat.ucla.edu/powercalc/normal/n-2/index.php.

Figure 4.2: Box and Whisker Plot Indicating Increased Compliance on C-Flex Compared to CPAP (Not Statistically Significant).



Removing patients who were unable to tolerate PAP therapy (i.e. those averaging less than 1 hours per night compliance) from the analysis resulted in the loss of 2 patients from each treatment group. Both compliance distributions were approximately normal when tested with the Kolmogorov-Smirnov test (both $p>0.7$) and the variance was not different between the reduced groups ($p=0.25$). Patients using CPAP averaged 3.8 hours per night (SD 1.5) whilst those who used C-Flex averaged 5.9 hours per night (SD 2.0). This 2.1 hour difference was statistically significant ($t=2.4$, 95% CI 0.2 to 4.1 hours, $p=0.035$, see Figure 4.3). The effect size was large (1.2), partially due to the reduction in variability. Using the same variability as the primary analysis (SD=2.5) still gives a large effect at 0.84. No other compliance cut-points were tested.

Figure 4.3: Box and Whisker Plot Indicating Statistically Significant Increased Compliance Levels on C-Flex compared to CPAP and when PAP Failures are Excluded.



Secondary Outcomes: Effects of Treatment on Sleepiness and Psychomotor Performance

All secondary outcome measurements were deemed approximately normal after testing with the Kolmogorov-Smirnov test (all $p>0.15$), and had approximately equal variance (all comparisons $p>0.5$). Independent t-tests indicated that the effects of treatment on secondary outcomes were not significantly different from one another, except for the Epworth Sleepiness Scale (see Table 4.5). Patients using CPAP were observed to have an 8.1 point mean improvement in the Epworth Scale (95% CI 4 to 12 points, $p=0.002$), compared with only 2.1 on C-Flex (95% CI -1 to 5 points, $p>0.05$). When compared to the baseline standard deviation (4.1 points, see Table 2), this 6-point net improvement gives an effect size of 1.46, a large effect. Neither the MWT nor the

PVT measures exhibited significant differences between the treatments. Both tended toward moderate improvements within the treatment arms, as shown by effect sizes. These were not statistically significant as this pilot trial was only adequately powered to detect differences of 1.3 standard deviations or more (i.e. Effect Size>1.3). Visual inspection of improvements in the PVT Mean Reaction Time and the PVT Lapses indicated some non-normality due to outliers, so these measures were additionally tested with the Mann Whitney U. These confirmed the findings of the t-tests ($p>0.35$ in both cases).

Table 5: Effects of Treatments on Secondary Outcome Measures

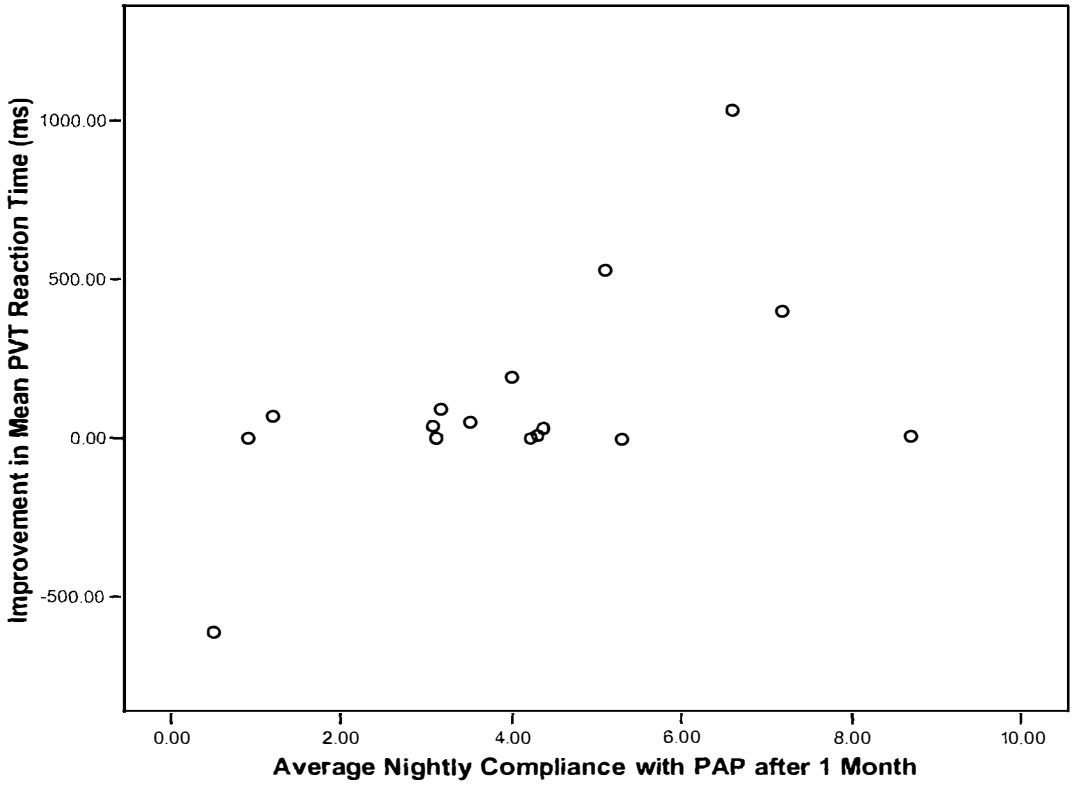
Outcome Measure	CPAP			C-Flex			P for difference between treatment effects
	Before Treatment Mean (SE)	After Treatment Mean (SE)	Effect Size	Before Treatment Mean (SE)	After Treatment Mean (SE)	Effect Size	
Epworth Sleepiness Scale (points/24)	15.0 (1.6)	6.9 (1.6)**	1.98	14.0 (1.3)	11.9 (1.6)	0.51	**0.014**
MWT (mins/40)	18.4 (2.7)	21.7 (3.1)	0.37	18.9 (3.7)	26 (3.4)	0.70	0.503
Mean Reaction Time (ms)	453 (128)	274 (19)	0.49	547 (149)	472 (200)	0.13	0.457
Number of Lapses (RT>500ms)	12.6 (6.1)	1.4 (0.7)	0.55	18.1 (8.7)	5.6 (4.1)	0.66	0.944

Stars indicate significant improvements from baseline * $p<0.05$, ** $p<0.01$. The direction and significance of these results were unaffected when the very low compliers with treatment were removed (mean <1 hour per night), as per the modification to the primary analysis. MWT=Maintenance of Wakefulness Test. RT=Reaction Time

Effects of Compliance on Secondary Outcome Measurements

The only significant correlation between compliance and secondary outcome measurement improvements was with PVT mean reaction times (r=0.552, p=0.03). However visual inspection of the scatter-gram, (see Figure 4) indicated that the strength of this correlation was mainly due to particular outliers and that the overall relationship was fairly flat.

Figure 4.4: Correlation Between use of PAP Devices and Improvements in Mean PVT Reaction Times After 1 Month of Treatment.

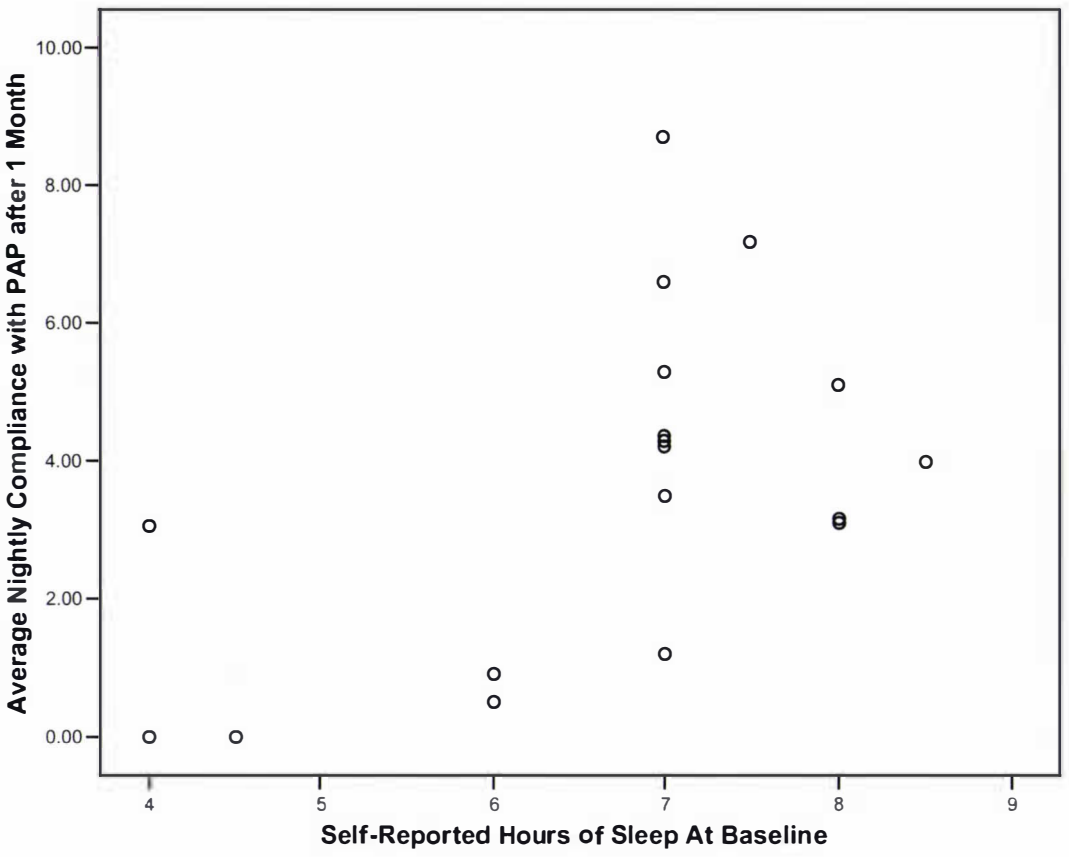


Predictors of Compliance

To identify possible pre-treatment indicators of compliance correlation analyses, were used to examine it's relationships with self-reported sleep hours, AHI, Arousal Index, sleep efficiency, BMI, CPAP pressure, baseline MWT latency, baseline PVT

mean score, baseline PVT lapses, and the baseline Epworth Sleepiness Scale. Only self-reported hours of sleep was significantly correlated to average compliance after 1 month (see Figure 5, Pearson's r=0.542, p=0.02).

Figure 4.5: Correlation between Self-Reported Sleep Duration and Subsequent Compliance with PAP therapy



The general shape of the relationship in Figure 5 does not lend itself well to a dichotomous decision rule for pre-treatment probability of high vs. low compliance. Patients with more than 6.5 hours per night of self-reported sleep still have a wide range of compliance.

4.4 DISCUSSION

This pilot randomised controlled trial comparing compliance levels in patients using standard CPAP with patients using C-Flex indicated that C-Flex may result in an extra 1.7 hours of compliance per night (95% CI -0.7 to 4.1 hours). Patients who were able to tolerate PAP therapy may gain an additional 2.1 hours of compliance (95% CI 0.2 to 4.1 hours/night) when using C-Flex compared to CPAP. This second group of patients is arguably a more relevant clinical grouping, giving the large proportion of patients unable to tolerate PAP therapy[Engleman, 2003 #412], possibly regardless of various subtleties in it's application.

However, this additional compliance did not result in better outcomes for patients. Epworth sleepiness scale improvements in the CPAP treated group were extremely large (mean 8.1 points), whilst they were moderate in the C-Flex group (mean 2.1 points). This difference in sleepiness reduction was statistically significant. Improvements in mean reaction times, lapses, and maintenance of wakefulness test latencies were usually moderate, and were similar between treatments.

Compliance did not seem to be meaningfully relate to improved patient outcomes. Nor was it possible, from data available at the beginning of the trial, to predict subsequent compliance.

4.5 CONCLUSIONS

A pilot randomised controlled trial has demonstrated clear primae facie evidence of increased compliance with C-Flex over that shown with CPAP. A larger trial is now warranted.

CHAPTER 5
CONCLUSIONS AND RECOMMENDATIONS

This thesis gives three examples of the potential information gains achieved by employing an evidence-based approach in sleep medicine. It applies two of the most powerful tools in EBM, the Randomised Controlled Trial (RCT) and the Systematic Review and Meta-Analysis of RCTs. This is a relatively new approach, with sleep medicine having been mostly developed using physiological experimentation, with reliance on polysomnography.

The findings from three EBM-based studies are discussed and placed in the context of the existing literature in the following sections.

5.1 RANDOMISED CONTROLLED TRIAL OF CPAP IN PATIENTS WITH MILD-MODERATE OBSTRUCTIVE SLEEP APNOEA

The use of CPAP to treat mild-moderate OSAS is potentially an example of clinical drift in sleep medicine. The use of CPAP in patients with severe OSAS has been shown to be effective at reducing sleepiness[149]. However, the 6 previously published RCTs involving only patients with mild-moderate OSAS, have produced conflicting findings over whether CPAP was effective. More data were clearly required at the time the present trial was funded in 1999/2000, when the findings from only 3 trials were available.

The present study found that, in patients with mild-moderate OSAS, humidified CPAP reduced subjective sleepiness, compared to humidified sham CPAP. There was a trend for objective wakefulness to improve, but this effect did not reach significance and was mostly due to worsening on placebo. CPAP did not improve any measures of reaction time, mood, or quality of life.

Study strengths

Significant strengths of the study included the use of humidification as an adjuvant to CPAP, the low level of dropouts (6.5%), relatively high compliance (4.9 hrs/night), and the novel use of sham CPAP in a crossover trial. Both the compliance levels and the dropout rate are superior to any other published RCT in mild-to-moderate OSAS.

Monasterio and colleagues[186] reported a 4.8 hours/night mean compliance with CPAP, but other trials, such as those reported by Barnes *et al.*[90, 171] (ca. 3.8 hrs/night) and Engleman *et al.*[184, 185] (ca. 2.8 hrs/night) report compliance levels considerably lower than in the present study. The higher compliance level in the present study might have been partly due to the addition of humidification to CPAP, which has been shown to marginally increase compliance[39, 178].

Post-randomisation dropout levels in previous studies have generally been between 8-13%[184, 187], but have also been reported as high as 30-33%[90, 171]. Our dropout rate was about 6.5%. However, a single extra participant dropping out of our study and/or one less from some other studies would have made our dropout rate undistinguishable from 4 of the 6 previous trials.

Chapter 2 was the first study to demonstrate that a CPAP device set at <1cmH2O (i.e. Sham CPAP) can be an effective placebo in a crossover design. There was a non-significant trend for patients to prefer placebo (17/29) and think it better for their sleep (19/29). Despite the published concerns[170], and an observation in a clinical trial[150] that placebo CPAP might have lower compliance rates than active CPAP, there was no significant difference in compliance rates in the present study. Nor were patients un-blinded, as feared[213]. Unlike other studies using a low pressured CPAP device[157, 162], the device used in the present study did not improve the major indices of sleep disordered breathing. Despite some worsening of overall sleep quality on sham CPAP, seventeen of twenty daytime outcome measures showed some improvement, addressing worries that the device might bias in favour of positive findings[213]. It is possible that the 3.1 minute worsening of ModMWT latencies on sham CPAP (p=0.15) was due to poorer sleep quality caused by sham CPAP. However,

a similar-sized, 2.7 minute negative effect has been observed when using a pill placebo[90]. These observations lead to a conclusion that the positive placebo effects were due to the expectation of benefit by the patients, and not to a partially beneficial or detrimental effect on sleep indices of the device. The negative placebo effects on MWT latencies are harder to explain. Statistically significant placebo effects were observed in a number of the subjective outcome measures including the FOSQ (2 of 4 subscales), the SF-36 (3 of 8 subscales), and the HADS (both subscales).

The placebo effects of sham CPAP and pill placebo seem to be similar. In Barnes' 2004 study, 8 of 14 outcomes (not including PSG and blood pressure indices) demonstrated placebo effects[90]. In the present study, 7 of 20 measures showed significant placebo effects. More would have been significant with a larger sample size. The magnitude of the placebo effects seems to have been similar on Epworth Scores (0.5 points here and in Barnes 2004, but 2.1 points in Barnes 2002[171]), MWT (3.1 minute worsening here and a 2.7 minute worsening in Barnes 2004), and FOSQ scores (improved by about 0.2 points both here and in Barnes 2004, and by about 0.3 in 4 of 5 subscales in Barnes, 2002). However, in our study a 6.5 point mean placebo improvement across SF-36 subscales would more than have accounted for the 4.7 point improvement on the CPAP arm in Barnes' 2004 data (about 8.6 in Barnes 2002). Nevertheless, overall, there seems to be little evidence that a pill placebo is inferior to sham CPAP, as some commentators have claimed[12, 168]. However, a direct comparison in an RCT would be needed to empirically resolve this issue.

Potential Study Weaknesses

Potential explanations for the lack of significant improvements due to CPAP therapy in the majority of outcome measures include: small numbers of participants (Type II error), lack of initial impairment, sub-optimal compliance, insufficient treatment duration, or that CPAP is not an effective treatment for patients with mild-moderate OSAS.

Type II errors were a potential problem with this study. However, post-hoc power calculations indicated that only two of the nineteen non-significant outcome measures were potentially under-powered. The ModMWT would have required 38 patients, and the FOSQ social outcomes scale 64 patients, to show differences that were significant. All other outcomes either favoured placebo over CPAP (n=10), or would have required more than 340 patients (three required many thousands) to show significance. The vast numbers of people required to show these tiny differences lends strength to a conclusion that the effects observed are largely indistinguishable from zero, and are thus not evidence of Type II errors. Another way to calculate whether our findings are subject to Type II error is to decide that effect sizes need to be at least moderate (i.e. 0.5) in order to be clinically meaningful. The main results table in Chapter 2 clearly indicates that only one outcome measure improved by such a margin. Ten of 20 outcome measures had negative effect sizes, because CPAP had been outperformed by sham CPAP. A simple power calculation indicates that, in order to show a moderate effect size (0.50) in any outcome measure with 80% power and a two sided alpha of 0.05 would require 33.4 patients. With 29 patients it should be possible to detect 80% of cases where the effect size is 0.539. Clearly our trial is not seriously

underpowered to detect effects of a moderate size. This trial confirms that the effects of CPAP, for mild-moderate OSAS, are clearly not as substantial as was once thought.

Twenty statistical hypothesis tests were run to test for the efficacy of CPAP therapy. The observed improvements in subjective sleepiness might have been a type I error (false positive). But, as the improvement fitted within an hypothesised dose-response relationship, and is consistent with the known effects of CPAP in severe OSAS[149], the result is likely to be both a real effect and clinically relevant. The effect sizes for the Epworth Sleepiness Scale (0.58) and the modified Maintenance of Wakefulness Test (0.40) were small to moderate, indicating potential clinical significance. However, ModMWT latencies tended to get worse by about 3 minutes on the placebo arm, generating more than half of this effect. The effect seen in clinical practice would be less than 0.20 (i.e. insignificant).

Our patient group exhibited sleep-related morbidity that was potentially amenable to treatment, with significant decrements compared to the population norms for the Epworth Sleepiness Scale (12.4 at baseline compared to 6.5 in the population[11, 193]); the Maintenance of Wakefulness Test (20.9 minutes at baseline compared to 35.2 or 36.9 minutes in the population)[195, 196]; in some subscales of the SF-36, notably Vitality (see Figure 5.1: Normative data adapted from Scott *et al.*[204]); and in some subscales of the Functional Outcomes of Sleepiness (see Figure 5.2: normative data adapted from Weaver *et al.*[198]). Figure 5.1 and Figure 5.2 show the clear deficits in patients with mild-moderate OSAS in the two health-related quality of life measures when compared to population estimates. They also show that significant morbidity remains in most subscales, after CPAP treatment. The subscales shown in the SF-36 are abbreviations of the following domains: Physical Functioning, Role Physical, Bodily

Pain, General Health, Vitality, Social Functioning, Role Emotional, and Mental Health. The subscales used in the Functional Outcomes of Sleepiness Scale are similarly as follows: Activity Level, Social Outcomes, Vigilance, and General Productivity.

Figure 5.1: SF-36 Scores in Mild-Moderate OSAS: Baseline Morbidity and Treatment Effects Compared to New Zealand Population Norms.

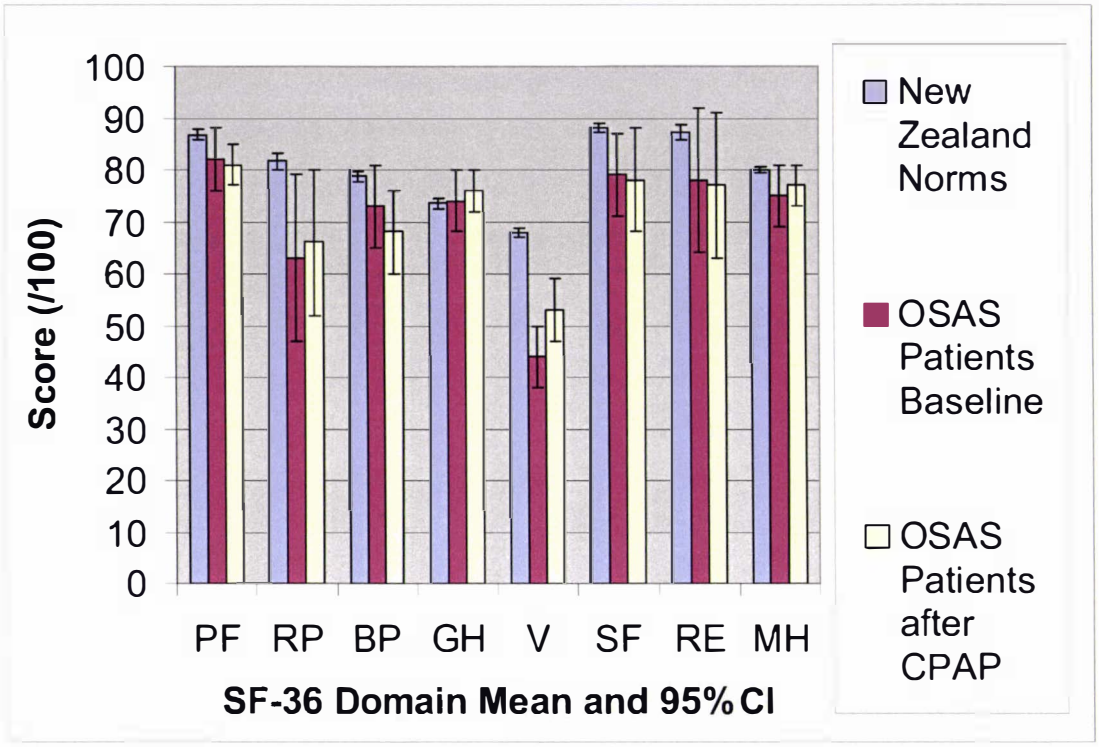
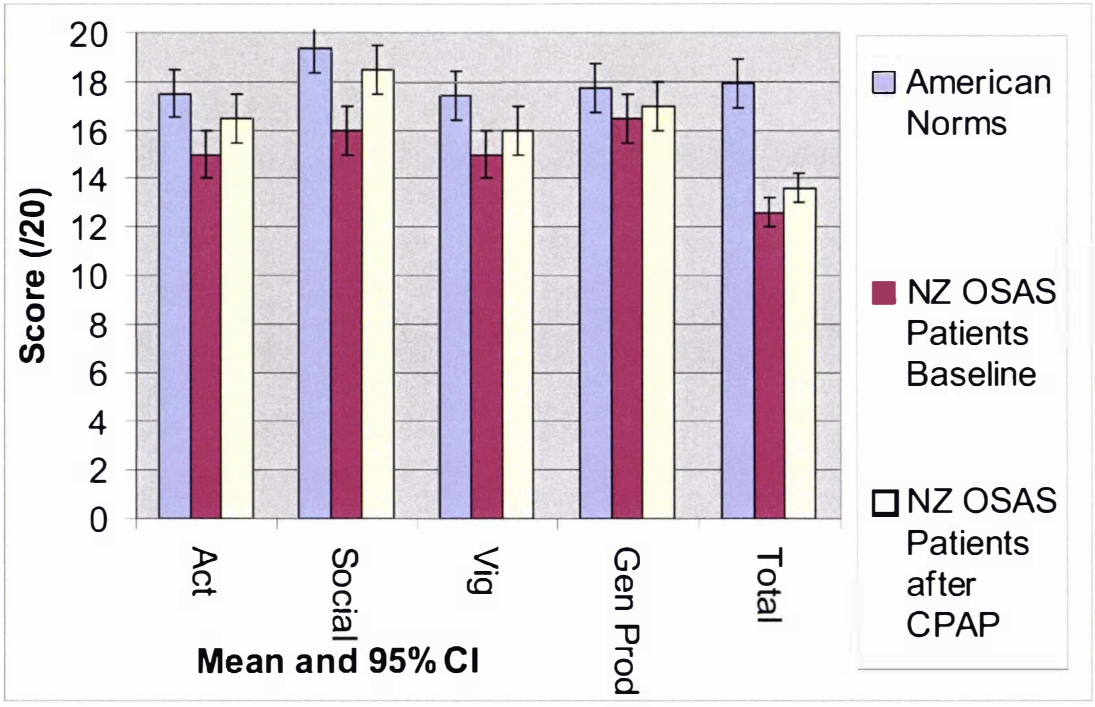


Figure 5.2: Functional Outcomes of Sleepiness Questionnaire in Mild-Moderate OSAS: Baseline Morbidity and Treatment Responses Compared to Population Norms.



Treatment durations in previous studies of mild-moderate OSAS have ranged from 4 weeks to 6 months, whilst this study used 3-week arms. The two trials[184, 185] employing 4-week treatment durations showed both treatment and placebo effects, whilst trials of only one week of CPAP in patients with a broad spectrum of OSAS severity have also shown benefits to mood[162], neuropsychological function[157], and even blood pressure[82]. The onset of CPAP-induced reduction in sleepiness occurs within a single night of treatment, and is further improved at 2 weeks, but not after 6 weeks[189]. In the present study, three weeks of treatment resulted in improvements in sleepiness, in addition to a number of placebo effects in mood and quality of life measures, indicating sensitivity over this period. Our treatment duration seems adequate for the outcome measures we have employed.

Sub-Analyses

As this was the first crossover trial employing a compliance-monitored sham CPAP device, it is possible to quantify some important effects. A hypothetical, physiological dose-response relationship would be confirmed if highly compliant CPAP users exhibited improvements that were better than both high and low placebo compliance groups. Only the Epworth Sleepiness Scale exhibited this expected pattern. Patients were not blinded as to their own compliance levels and were thus able to infer some of the benefit they might expect, independent of the objective benefit CPAP might confer. Previous studies in this field have found that high users of CPAP tend to show greater improvements than low users of CPAP[184, 185, 187]. This might be in part due to their expectation of benefit, which could not be quantified by compliance level due to the employment of a pill placebo. Our trial provided the first supporting evidence for

this problem from a crossover trial. Patients who were high users on CPAP did well, but in all cases, except the modMWT and Epworth scales, they did not do better than when they received sham CPAP. This suggests the benefits of the high use of CPAP are partially due to the knowledge of high use, and potentially other important non-randomised factor/s. This phenomenon presents an issue to clinicians using such measures to monitor the progress of individual patients using CPAP. High compliance effects cannot be separated into the true benefits of CPAP and the psychological expectation of benefit that occurs with high levels of use.

Sub-analyses also showed that the most likely group to enjoy reductions in Epworth scores were those with high Epworth scores at the outset. This supports Engleman's recommendation that the sleepiest patients with mild OSAS should be treated with CPAP[210]. However, given that there is confusion between the selection criteria and the main outcome of interest (we have used Epworth to screen for Epworth reductions), it might be expected that substantial reduction to the mean might account for this phenomenon. Fortunately in this study, we were probably able to control for this by subtracting the changes during the placebo arm from those in the CPAP arm. This should control for the placebo effect/s and should also control for any natural tendency for regression to the mean. It is also possible that, in patients with low Epworth scores (<11), there might be a floor effect whereby Epworth scores could not improve by enough to make a large effect (i.e. 4 points or more). Conversely in those with high Epworth (>15), it might be less likely that scores would worsen. Nevertheless, the present study does suggest that patients with higher initial Epworth Scores are more likely to show significant improvement in sleepiness as a result of CPAP treatment.

Recommendations for Further Research

This particular concept deserves specific testing in a placebo controlled RCT. A suitable trial would require patients with Epworth scores of at least 16, mild-moderate SDB (AHI 5-30), and could not use the Epworth as an outcome measure. The recommended primary outcome measure is the Maintenance of Wakefulness Test, given that it seems to be more sensitive to treatment than the Multiple Sleep Latency Test[153].

Because of the significant, quantifiable effects of real and placebo compliance, it is recommended that future placebo controlled studies of CPAP use a sham CPAP as the placebo, where possible.

Advantages of sham CPAP as a placebo are:

1. Effective control of the negative effects on sleep of CPAP;
2. Requiring the same amount of effort and active participation on the part of the patient as does CPAP;
3. The ability to objectively quantify compliance (possibly covertly).

Sham CPAP has been demonstrated to be safe[169] and has now been used in a number of RCTs[79, 83, 150, 155, 156, 164, 244].

The present study has shown that sham CPAP can be used in a crossover trial. Direct comparison between compliance of parallel groups has been useful in previous trials, but the direct correlations within patients of compliance on placebo and CPAP presents the most convincing evidence that compliance was stable between treatments.

The comparison of the effects of compliance within patients, between active or inactive treatments, is also instructive. Although some study aims preclude the use of crossover trials, this study design is recommended wherever possible to reduce the numbers of people subjected to randomised treatment.

The weakness of this approach is that the compliance effect itself is a non-randomised effect, but this is a problem common to many RCTs, not just those of CPAP. Randomising the dose of CPAP in further trials, much as is done in pharmaceutical trials, might shed light on the magnitude of this potential effect in CPAP research. The 'dose' in this case would need to be time at effective pressure. The difficulty would be designing and operating a device that can switch itself off, and being able to persuade low compliant patients to use the machine to the required dose. The same difficulty is evident in the sleep restriction literature, where individuals in randomised groups sometimes do not sleep the full amount their group has been allotted[102, 103].

Mild-Moderate OSAS CPAP Trial Conclusions

CPAP moderately reduced subjective sleepiness, and may have improved objective wakefulness, in patients with mild-moderate OSAS. It did not improve reaction times, mood or health-related quality of life. The findings of the present study, when viewed in context, offer little evidence that CPAP should be routinely used in the treatment of mild-moderate OSAS, but offer provisional support for the practice of trialling CPAP in patients with mild-moderate OSAS and severe daytime sleepiness[210]. The study also offered some important methodological to consider in reviewing the overall literature. Further trials are needed to assess the potential to control cardiovascular disease in these patients.

5.2 SYSTEMATIC REVIEW AND META ANALYSIS OF RANDOMISED CONTROLLED TRIALS OF CPAP FOR TREATING DAYTIME SLEEPINESS IN PATIENTS WITH MILD-MODERATE OBSTRUCTIVE SLEEP APNOEA SYNDROME.

The previous systematic review and meta-analysis of the effectiveness of CPAP was unable to conclude that CPAP was effective in reducing sleepiness in a sub-analyses for mild-moderate OSAS[149]. The effect seemed to be slight, if not negligible. Even if the observed effect was real, the combined studies were underpowered to detect it. One study was incorrectly misclassified as being of mild-moderate severity and may have diluted the real effect[83]. Since that time, a further three trials have been published. Another meta-analysis based on a systematic review of all randomised controlled trials, of the type reported in Chapter 2, was required.

Main Findings

The meta-analyses presented in Chapter 3 have shown that CPAP reduced subjective daytime sleepiness by 1.2 points on the Epworth Sleepiness Scale (95% CI 0.5 to 1.9, effect size 0.27) and improved objective daytime wakefulness measured by the Maintenance of Wakefulness Test by 2.1 minutes (95% CI 0.5 to 3.7, effect size 0.21), in patients with mild-moderate OSAS. However, the size of the two effects is small, and close to being clinically irrelevant (i.e. <0.20[112]). There was insufficient evidence to claim that CPAP decreased objective daytime sleepiness, as measured by the Multiple Sleep Latency Test, in these patients. MSLT latencies non-significantly worsened by a mean 0.2 minutes (95% CI -1.0 to 0.6).

The findings from the MWT analysis need to be interpreted with caution. Both of the trials that reported placebo effects noted worsening, which was at least as great as the benefit reported in the present meta-analysis. Barnes *et al.*[90] have noted a 2.7 minute, and Chapter 2 noted a 3.1 minute worsening during the placebo arm[38]. In 2004, Barnes *et al.* reported that the effect of CPAP alone (unadjusted for placebo) was for sleep latency to worsen by 0.7 minutes, whilst Chapter 2 reported a 2.1 minute improvement on CPAP. The estimate of effect shown in Chapter 3 is heavily reliant on the negative influence of placebo treatment on wakefulness. This unexpected pattern might indicate a real effect, where wakefulness is becoming subtly more difficult to maintain in these patients about the time they are invited into a clinical trial. However, the effect is more likely to be evidence of a slight test-retest worsening of MWT latencies in these crossover trials, possibly due to test dissatisfaction and boredom, or to increased comfort within the sleep laboratory. The meta-analysed effect is not a true indication of the effect that would be seen in clinical practice, which would be substantially smaller and possibly not different from zero. The possible test-retest worsening of MWT latencies merits further attention and should be monitored in any future RCT studies, by including comprehensive baseline measurements.

Sub-Analyses

Observed Epworth Sleepiness Scale heterogeneity did not seem to be moderated by gender, age, body mass, baseline Epworth, or the patients' objectively measured compliance with treatment. Nor did the effect seem to be modified by study characteristics, including trial quality (Jadad scores), crossover or parallel design, length

of follow-up, multi or single study venues, country of origin, or control comparison employed. Trials with low dropout rates did, however, show significantly better outcomes in terms of Epworth score. This seemed to be a chance finding with the effect being in the opposite direction to that expected. Neither the Multiple Sleep Latency Test nor the Maintenance of Wakefulness Test were found to have significant heterogeneity.

Sensitivity analyses of all three outcome measures did not indicate any studies which had undue influence upon the final estimated effects of CPAP. Funnel plots, and associated statistics, offered no proof that the effects observed were due to selective publication, but this was to be expected given the relatively small size of these trials. However, it is always possible that a number of unpublished trials exist.

Integration and Implications of Findings

These results extend the recent meta-analysis by Patel and colleagues[149]. The mean improvement in subjective sleepiness in their 'mild-moderate' studies, was a non-significant 1.1 points on the Epworth Sleepiness Scale, compared to a significant improvement of 4.75 points in the studies of (largely) severe OSAS. Our finding that subjective sleepiness improved by 1.2 points is almost identical in magnitude, but was statistically significant due to the additional power derived from including the three most recent studies[38, 90, 171]. The clinical applicability of our findings was also aided by the exclusion from the analyses of the study by Barbe and colleagues[83], which examined the effects of CPAP on patients with severe SDB, but no abnormal daytime sleepiness.

MWT latencies in the studies analysed by Patel *et al.* revealed a mean benefit of 3.0 minutes across the full spectrum of OSAS. Thus, our finding of a 2.1 minute improvement in MWT latencies in patients with mild-moderate OSAS, is concordant. Quantification of the magnitude of the effects, by using effect sizes, was a strength aimed to aid clinical interpretation of our findings.

Whilst Patel *et al.*[149] did not find significant differences between MSLT and MWT sensitivities to CPAP treatment, the present analyses were constructed with the assumption that these metrics should not be combined, as they probably measure different abilities (the ability to stay awake versus the ability to fall asleep)[152]. The separating of sleepiness and wakefulness also seems to be supported by findings from the neuroscience literature which indicated that these two processes arise from separate but interacting neurological substrates[245]. The 95% confidence intervals from the meta-analysed estimates indicated that whilst the MSLT scores were not systematically affected by CPAP therapy, the MWT scores were significantly improved. Furthermore, improvements in latencies on these two tests were significantly different ($p<0.05$).

It is not clear from the present analyses whether benefit accrues most to those patients with moderate OSAS (AHI 15-30), or equally to those across the mild-moderate spectrum, or whether there exists an average severity point at which CPAP is no longer effective. It is also unclear whether the expense associated with ongoing CPAP treatment is cost-effective for reducing daytime sleepiness in mild-moderate patients. This is particularly troublesome when the effect sizes are very small and close to being insignificant. It also remains to be demonstrated in an RCT whether the putative cardiovascular benefits of CPAP exist for patients with mild-moderate OSAS, or whether CPAP is a cost-effective treatment for improving cardiovascular disease risk

in this group.

It is conceivable that the findings of the 1100 person multi-centre APPLES trial in the USA, which is comparing the effects of CPAP and sham CPAP in two groups treated for 6 months, might reach different conclusions to this meta-analysis http://clinicaltrials.gov/ct/show/NCT00051363?order=3 . The APPLES trial has planned sub-group analyses in the mild-moderate spectrum and will probably be a higher standard of evidence due to the superior sample size, despite the generally accepted EBM hierarchy. The smaller, but similarly designed CATNAP study in men with mild OSAS (n=280) http://clinicaltrials.gov/ct/show/NCT00089752?order=1 might also offer a differing interpretation of the efficacy of CPAP therapy.

Mild-Moderate OSAS CPAP Meta-Analysis Study Conclusions

Continuous Positive Airway Pressure significantly improves subjective sleepiness and objective wakefulness in patients with mild-moderate obstructive sleep apnoea. It does not improve objective sleepiness. However, the effect sizes of these statistically significant adjustments are very small and might not be clinically relevant or a cost-effective use of limited sleep medicine resources.

5.3 PILOT RANDOMISED CONTROLLED TRIAL OF A NOVEL SELF-MODULATING CPAP (C-FLEX) COMPARED TO STANDARD CPAP.

Technical drift is a potential problem inherent in the technological race to improve CPAP comfort. New adjuncts to therapy may not deliver the benefits they advertise, or the benefits might be so slight as to not be cost-effective. C-Flex adjunct to CPAP reduces the pressure during exhalation, presumably to remove some of the extra effort required to breath out against positive pressure. The manufacturer claims that C-Flex is more comfortable than CPAP, resulting in higher compliance. We tested in a pilot RCT whether C-Flex resulted in higher compliance than CPAP in patients with severe OSAS. A crossover trial would have been a superior choice to our parallel design. However, we undertook this study without external funding and were largely constrained by this into running our trial within existing clinical channels.

Primary Analyses: Effects on Compliance

Patients randomised to C-Flex had 1.7 hours/night higher compliance than those randomised to CPAP (95% CI -0.7 to 4.1 hours). This difference was not statistically significant due to under-powering ($p=0.15$), but was of moderate size (Effect Size =0.68). By comparison, the effect of the commonly-used heated humidification adjuvant is about 0.25-0.30[39, 178, 225].

However, around 5-50% of patients prescribed PAP therapy are unable to tolerate the device and abandon treatment within the first week after prescription[225]. When these non-tolerant patients (defined as less than 1 hours compliance per night on

average, n=4/19) were removed, the difference in mean compliance between C-Flex and CPAP increased to 2.1 hours, became statistically significant (95% CI 0.2 to 4.1 hours, p=0.035), and large (conservatively ES=0.84).

In June 2005, a 3-month non-randomised study, indicated that moderate-to-severe OSAS patients who received C-Flex had on average, 1.7 hours/night higher compliance, than those receiving CPAP[238]. This margin is identical to the one identified by our study and the two groups exhibited very similar absolute compliance levels. The effects on secondary outcome measures (FOSQ and Epworth) were similar between treatments. That paper also contained reference to an abstract[246] reporting the findings of a small RCT (n=16) of C-Flex compared to CPAP. Loube and Ball found that C-Flex resulted in a non-significant increase of 0.6 hours/night (36 mins, p=0.09). There was no difference between treatments in improvements on the FOSQ or the Epworth scale.

Secondary Analyses: Effects on Daytime Function

Increased levels of compliance with C-Flex did not result in enhanced improvements in the secondary outcome measures. Indeed CPAP therapy caused a significantly greater reduction in the Epworth Sleepiness Scale than did C-Flex (8.1 points vs. 2.1 points, p for diff=0.014). This 6 point difference in Epworth is also clinically significant, with a large effect size of 1.46. The mean post-treatment Epworth Score in the CPAP group was 6.9, placing it about the same level as seen in the general population in New Zealand[11]. However the C-Flex group was still significantly somnolent, at a mean of 11.9 points.

Other secondary outcome measures (Modified Maintenance of Wakefulness Test, Mean PVT Reaction Time, and Lapses on the PVT) tended to improve by about the same amount in both treatments, but these improvements were not statistically significant. However, the improvements tended to be moderate in magnitude (ES 0.49-0.79). For instance, improvements on the MWT in this study (3.3 mins on CPAP and 7.1 mins on C-Flex), were broadly similar to those seen in CPAP RCTs. The most recent meta-analysis of CPAP RCTs for severe OSAS found a 3.0 minute improvement in MWT latencies[149]. Normative values for the MWT indicate that a mean sleep latency in a healthy group should be above 30 minutes[195, 196]. In the present study, mean latencies were at 22 and 26 mins (CPAP and C-Flex respectively) after treatment. Thus, some residual daytime sleepiness remained in both groups, despite the moderate treatment effects.

The PVT mean reaction time improved by 179ms on CPAP (from 453 to 274 ms) and 75ms on C-Flex (from 547 to 472 ms), whilst the numbers of lapses improved by 11.2 on CPAP and 12.5 on C-Flex (12.6 to 1.4, and 18.1 to 5.6 respectively). The reduction in mean reaction time on CPAP was small/moderate in magnitude, but statistically non-significant due to the small sample size. Nevertheless, CPAP did bring the mean RT into a range that might be regarded as relatively normal for middle-aged people[247]. Lapses (reaction times longer than 500ms) also improved by a moderate amount (12 points) during both treatments, but this was not statistically significant. While the gross size of the reduction was similar in both treatments, patients on C-Flex started with a higher average number of lapses and did not reduce to the relatively normal levels of those on CPAP.

Post-hoc Analyses

The limited polysomnographic evidence collected suggests that C-Flex treatment was probably as effective as CPAP at controlling SDB. However, the use of split-night studies limited the ability to determine whether the CPAP and C-Flex groups differed pre-treatment in terms of SDB indices or sleep architecture, or whether the treatments had different effects on sleep. Any differences might have explained the puzzling differences in ESS scores. This trial was only designed to examine whether a large difference in compliance might exist. A well-designed randomised polysomnographic investigation is needed to assure the sleep medicine community that C-Flex is as technically effective as CPAP for controlling SDB.

Correlations between compliance and reductions in sleepiness were calculated across treatments, due to lack of numbers in this small trial. However, despite the logical purported benefits of longer durations of treatment use, this trial showed no clear benefits to those who were highly compliant. Neither subjective nor objective improvements showed a typical dose-response relationship, despite the generally moderate to large effect sizes. Unfortunately, use of multivariate models to predict which patients might show improvement in secondary outcome measures, was not feasible in such a small sample.

The relationship between shorter hours of sleep and decreasing compliance might indicate that those with sleep disordered breathing and insomnia are poor candidates for PAP therapy. Or it might indicate that in these patients daytime sleepiness is caused by lack of sleep rather than SDB. Poor compliance then results because PAP offers no symptomatic relief, which is the strongest correlate of CPAP compliance[225]. Patients might be unwilling or unable to use PAP due to the correlates of insomnia, including anxiety, claustrophobia, or depression. However, due to the post-hoc nature of the test, this finding needs to be interpreted with extreme caution.

Recommendations for Further Research Trials Arising from the C-Flex Study

Proposed Randomised Trial of Compliance on C-Flex Compared to CPAP

This pilot study had some significant, but somewhat contradictory findings, indicating a clear need for a larger study. The variability in compliance amongst C-Flex users seems to be comparable to that observed amongst the users of CPAP. As such, power calculations for future trials can assume that C-Flex will match local CPAP compliance variability.

Based on the findings here detecting a, parsimonious, 1-hour difference in compliance would require 99 patients per group (given a common standard deviation of 2.5 hours). In a crossover trial (where a washout must be included) 38-64 patients would be required to show a 1-hour difference between C-Flex and CPAP compliance, given intra-individual variability of 1.5-2.0 hours per night. A key advantage to a crossover design, other than reducing the numbers needed to show this difference, would be to gain blinded treatment preference data from the patients. Such a trial, or trials, could then be combined in a meta-analysis with the present trial, to give a more accurate reflection of the effectiveness of C-Flex at increasing compliance and reducing sleepiness, compared to CPAP.

The present trial was relatively short, at 4 weeks. Further trials may need to be longer, depending on the sensitivity of the outcomes that are to be investigated. In this particular case, the length of trial was sufficient to show effects on compliance and on trait subjective sleepiness. Four weeks may not be sufficient for other outcomes, particularly markers of cardiovascular disease risk.

Ideally, long-term treatment compliance levels should be shown to reduce either cardiovascular disease, or well established markers of CVD risk, compared to placebo. It would be possible to investigate carefully selected CVD markers in a better-powered and longer RCT. If such a trial were to find that C-Flex had higher levels of compliance than CPAP, and that C-Flex reduced marker/s of CVD more than CPAP, and that compliance was correlated with reductions in CVD risk factors, then C-Flex might replace CPAP as the gold standard for treatment of OSAS (assuming that both of these

devices are better than placebo). Even then, further cost-benefit analyses would need to show that a move to C-Flex was worthwhile.

Proposed Randomised Trial of SDB Control with C-Flex Compared to CPAP

Because we were unable to confirm the comparability of SDB control between C-Flex and CPAP with any precision, a well-controlled polysomnographic study is recommended. This would require the separation of the diagnostic and titration phases of polysomnography into separate nights, in order to provide better information on sleep cycles across the night with and without treatment.

An ideal study design would involve patients who are already highly compliant with CPAP therapy undergoing PAP re-titration on night one. Over the next two nights, blindly and in random order, they would then use C-Flex or CPAP at the newly-titrated pressure. PSG recordings should be scored by technicians and physicians who are blinded to treatment allocation or to the study aims. The effects of CPAP and C-Flex across a full night of sleep could then be compared, after controlling for any pressure requirement changes in the patient, sleep stage, sleep position, and for any first night or order effects. The use of highly compliant patients allows the maximisation of data on the effects of therapy, and also allows patients familiar with positive airway pressure treatment to state their preferred treatment under blinded conditions.

Powering such a trial requires an estimate of an acceptable margin of precision that C-Flex must approximate compared to CPAP, in order to determine equivalence of the two treatments in controlling SDB.

C-Flex Trial Conclusions

C-Flex seems to increase compliance when compared to CPAP (mean difference 1.7 hours/night 95% CI -0.7 to 4.1 hours/night). However, improved compliance did not result in superior patient outcomes for waking function. These findings require replication in larger proposed studies.

5.4 THESIS CONCLUSIONS

The three studies presented add significantly to the evidence-base in sleep medicine. Chapter 2 is the first sham CPAP controlled crossover trial which is an approach previously thought to be impossible. Chapter 3 is the first meta-analysis to show a significant measurable effect on any metric of daytime function in patients with mild-moderate OSAS, although these effects are small. Chapters 2 and 3 together indicate that the widespread use of CPAP to treat mild-moderate OSAS is probably an example of clinical drift as the evidence of benefits accruing to patients is not compelling. This conclusion should be overturned if CPAP is demonstrated to effectively reduce the excess cardiovascular risk observed in these patients. However, in the present capped funding environment to treat a patient with mild-moderate OSAS is to effectively deny efficacious treatment to a patient with severe OSAS. Chapter 4 is the first full description of a RCT of a new CPAP like treatment modality called C-Flex. The trend toward a substantial increase in compliance above that observed with CPAP suggests that this device is not an example of technological drift as we had originally surmised. However this needs to be confirmed by larger trials that are suggested.

These three studies all represent original substantive contributions to evidence-based sleep medicine and are all either published or are being reviewed for publication by leading international respiratory medical journals.

5.4.1 Clinical Drift and the Treatment of Mild-Moderate OSAS

Our randomised controlled crossover trial of humidified continuous positive airway pressure (CPAP), compared to sham CPAP, for treating mild-moderate obstructive sleep apnoea, showed few benefits for waking function. CPAP brought about improvements only in subjective sleepiness (Epworth Sleepiness Scale) and potentially in objective wakefulness (Maintenance of Wakefulness Test). There were no improvements in measures of mood, quality of life, or reaction times. It is therefore recommended that CPAP should not be used in all patients with mild-moderate OSAS. However, there did seem to be larger reductions in sleepiness in those patients with greatest daytime sleepiness at baseline.

Most comparable trials are relatively small and come from disparate geographical locations. The best method for aggregating high quality RCT information is with a systematic review and meta-analysis. Our systematic review identified seven randomised placebo or conservatively controlled trials of CPAP in patients with mild-moderate OSAS. When meta-analysed, CPAP improved subjective sleepiness (Epworth Scale) and objective wakefulness (Maintenance of Wakefulness Test), but not objective sleepiness (Multiple Sleep Latency Test). However, the two positive effects were very small and of questionable clinical significance. These findings reinforce the recommendation that not all patients with mild-moderate OSAS should be treated with CPAP, as it does not effectively treat their morbid levels of sleepiness.

There are a number of possible explanations as to why CPAP does not markedly reduce sleepiness in patients with mild-moderate OSAS.

First, sleepiness may not be due entirely to SDB in all patients. The amount of sleepiness reduction seen does not bring these patients into line with population norms for either the ESS or MWT[11, 196]. Poor compliance with therapy fails to account for this remaining untreated morbidity. The two RCTs reported here found poor correlations between compliance and relief of daytime symptoms. Furthermore, the heterogeneity in meta-analysed Epworth improvement was not explained by the average CPAP compliance extracted from each of the studies. In addition, CPAP has a residual half-life for the control of SDB after the device is taken off[248-250], which further dilutes the potential effect of partial compliance. Thus, a number of additional questions arise regarding the pathogenesis of sleepiness in these patients.

Epidemiological studies indicate that SDB(AHI>5) is common, as is daytime sleepiness. Sleep disordered breathing has been found to affect about 25% of middle-aged males and 9% of females[3, 4]. On the other hand, daytime sleepiness (Epworth >10) affects about 15% of the New Zealand middle-aged population[11] and about 15% of the participants who never snorerd from the US Sleep Heart Health Study[251]. There is only a weak relationship between the severity of SDB and daytime sleepiness in epidemiological studies[3, 53], as such many of the patients seen in a sleep clinic with mild-moderate OSAS could have the combination of SDB and daytime sleepiness co-existing by chance alone. The concept that these two relatively common phenomena might exist together non-causally is not new[12]. Furthermore, many individuals with severe SDB seem to be relatively un-sleepy[53], and there is a great inter-individual

variability in the effects of sleep restriction[103, 188, 252, 253]. Another possibility is that SDB and daytime sleepiness are both caused by a third factor, perhaps, as has recently been proposed, insulin resistance and/or low grade chronic inflammation[13].

Another potential explanation for the lack of treatment effect might be that repetitive oxygen desaturations over many years of SDB cause irreversible damage. Some proof of this has been found in the cortical imaging studies reviewed in Chapter 1. Macey and colleagues[117], for instance, found diffuse bilateral damage to the cortex, which was consistent with a pattern of damage caused by oxygen deprivation. There is also some recent murine support for this hypothesis , with prolonged intermittent hypoxia shown to cause hyper-somnolence and lesions in the sleep/wake brain system[254].

The lack of an observed treatment effect might also be due to the short duration of the trials. Given that SDB is likely to have a progressive effect on daytime functioning, longer trials might be necessary to establish whether CPAP halts, or ameliorates, a decline in function. The longest duration RCTs available in severe OSAS lasted 1 year[255, 256] and measured sleepiness at 3 months and at 12 months via a visual analogue scale. It did not find systematic worsening in the control group. Neuropsychological functioning in controls tended to improve over the testing period, which was consistent with commonly observed practice effects. Another study[223] showed no significant worsening of sleepiness over a sixth month period of conservative management. It is possible that 6-12 months might be an insufficient time for CPAP to ameliorate a deterioration in function due to OSAS. Long term data from the Sleep Heart Health Study should help clarify the natural progression of SDB.

CPAP might justifiably be used as a treatment option for mild-moderate OSAS, if it is demonstrated to reduce CVD risk. Data being collected from the Wisconsin and Sleep Heart Health Study cohorts indicate that increased risk of cardiovascular disease exists even in what has been traditionally regarded as mild SDB, and that as the severity of SDB increases, so too does the risk of CVD[7]. SDB is prevalent in New Zealand, with Māori[33], and probably Pacific Islanders[34], particularly at risk. Large RCTs of CPAP effectiveness in patients with severe, moderate and mild OSAS are needed, that are suitably designed and powered to measure possible changes in cardiovascular disease risk (or established markers), over periods of at least a year or longer.

A further complication is that the vast majority of people with SDB do not have daytime sleepiness, and therefore do not tend come to the attention of clinical sleep services. It seems highly unlikely the large proportion of the population with asymptomatic SDB can be convinced to wear cumbersome CPAP to theoretically reduce their CVD risk, whilst enjoying no day-to-day benefits.

CPAP may not be the best treatment option for reducing CVD risk. For instance, two recent trials have indicated that mandibular advancement splints can significantly reduce some 24-hour blood pressure indices in patients with mild-moderate OSAS[89, 90], whereas CPAP did not[90]. Standard pharmacological agents might also be useful, but these have yet to be tested specifically in people with SDB.

New treatment options other than CPAP, need to be developed for mild-moderate OSAS, and rigorously tested by independent clinical researchers in randomised placebo controlled studies.

Implications for Treatment Services

The studies in Chapters 2 and 3 clearly show that clinical drift has occurred. The widespread treatment of mild-moderate OSAS with CPAP is without any marked benefits to patients' daytime sleepiness. The Australasian treatment guidelines released by the National Health and Medical Research Council of Australia[111], whilst attempting to address significant need, are clearly not evidence-based guidelines. The present meta-analyses (Chapter 3) show that many patients who would be treated under these guidelines would not enjoy large relief of their daytime sleepiness after treatment with CPAP. The guidelines should be modified to take account the evidence presented here, and the also the findings from the APPLES and CATNAP studies, when they become available.

There also remains a need for more trials that specifically include woman, ethnically diverse populations, and the elderly (aged 65+). Data regarding the effectiveness of CPAP for treating these groups is lacking for all severities of OSAS. Currently only one study has been powered specifically to examine CPAP effectiveness for treating OSAS in women[187].

As the implementation of CPAP therapy costs the same regardless of severity, and the greater benefit is derived from treating groups with severe OSAS, it is clearly more cost-effective to treat the severe group preferentially. Where there is a large untreated population and limited resources with which to treat them, those most effectively treated should be treated first. This logic must be applied to rationally planned socialised medical systems such as New Zealand, Australia and the United

Kingdom. In such a system, to ineffectively treat mild-moderate OSAS is to deny effective treatment to those with severe OSAS.

The lack of clinical services in New Zealand makes treatment of even severe symptomatic OSAS difficult. Treatment of severe OSAS has been shown to largely ameliorate a high level of observed mortality (15-20% over 5 years), observed in a clinical cohort[100, 257]. This potential is being missed in New Zealand. Investment in the development of an appropriate level of clinical services is required, otherwise any research efforts are wasted.

Technical Drift in OSAS Treatment

Technical drift might be occurring with the marketing of the new C-Flex adjunct to CPAP therapy, which aims to increase patients comfort and compliance by modulating pressure during exhalation. Our pilot RCT showed that patients using C-Flex were possibly more compliant than patients using standard CPAP and that the size of this effect was moderate to large. However, even in such a small trial as this, CPAP was significantly better at reducing daytime sleepiness as measured by the Epworth Sleepiness Scale than was C-Flex. Increased compliance with C-Flex failed to elicit additional daytime benefits to patients.

Given that considerable resources are invested in increasing compliance with CPAP therapy, via the introduction of new technological advancements, should we not have a clear idea of what we are buying? There are few studies that investigate dose-

response (treatment duration, not CPAP pressure) curves from CPAP therapy[248]. It is not known what additional relief, if any, is gained by even moderate increases in compliance. Evidence was presented in Chapter 2 suggesting that part of any benefits in waking function observed with high compliance are potentially psychosomatic (belief that being compliant with treatment will improve outcomes).

Some commentators have implied that the range in compliance is an individually optimal dosing strategy[228, 248-250, 258]. As patients can control the amount of use of a CPAP device much more accurately than when using pharmaceutical treatments, they are able to control their duration of CPAP therapy. Patients might be self-titrating to trade-off between maximising subjective benefits whilst minimising mask time. This could be reinforced by the beneficial effects of CPAP in reducing AHI and oxygen desaturations when the mask is not worn for the second part of the night, presumably due to increasing airway calibre and/or reducing pharyngeal oedema[249, 250].

Together, these factors imply that there could potentially be a marginal decrease in the benefit accruing from CPAP treatment as the hours of CPAP treatment per night increase. If patients are using this sort of strategy, it might explain why the hypothesised dose-response relationships are so poorly supported by the data. If the level of trade-off between use and daytime benefit, or use and residual night-time relief of SDB without wearing CPAP, is highly variable between individuals, then it would explain why patients who give themselves two hours of CPAP per night can see as much gain as a person who uses CPAP for 8 hours per night. It might also explain why patterns such as in Chapter 4 are seen, where outcome improvements are evident, but in the mid-ranges of compliance, benefits seem to have little to with the duration of use.

Chapter 2 shows that the compliance levels between treatments seen within each patient were highly stable, despite the marked pressure difference between a sham CPAP device (pressure <1cmH2O) and standard titrated CPAP (4-10cmH2O). This suggests that compliance levels may have little to do with applied pressure levels. One recent meta-analysis[259] found that moderate pressure reduction by AutoPAPs (mean 2.2 cmH2O) did not result in an increase in compliance. One the other hand, the findings in Chapter 4 may provide some counter-evidence, since C-Flex did result in increased compliance. The new C-Flex device clearly warrants further investigation in a properly powered independently-conducted RCT. However, it should be noted that the benefits of marginal increases in compliance are of uncertain benefit.

References

1. Dement W and Vaughan C, *The promise of sleep*. 1999, New York: Dell Publishing.

2. American Sleep Disorders Association, *International classification of sleep disorders (rev.)*, ed. MJ Thorpy. 1997, Rochester: ASDA.

3. Young T, Palta M, Dempsey J, Skatrud J, Weber S, and Badr S, The occurrence of sleep disordered breathing among middle aged adults. *NEJM*, 1993. **328**: 1230-35.

4. Bearpark H, Elliot L, Grunstein R, Cullen S, Schneider H, Althus W, and Sullivan C, Snoring and sleep apnea: A population study in Australian men. *Am J Respir Crit Care Med*, 1995. **151**(5): 1459-65.

5. Lavie P, Herer P, and Hoffstein V, Obstructive sleep apnoea syndrome as a risk factor for hypertension: population study. *BMJ*, 2000. **320**: 479-82.

6. Stradling J, Pepperell J, and Davies R, Sleep apnoea and hypertension: proof at last? *Thorax*, 2001. **56**(9): 45-60.

7. Young T, Peppard PE, and Gottlieb DJ, Epidemiology of Obstructive Sleep Apnea: A Population Health Perspective. *Am J Respir Crit Care Med*, 2002. **165**(9): 1217-39.

8. Lindberg E and Gislason T, Epidemiology of sleep related obstructive breathing. *Sleep Med Rev*, 2000. **4**(5): 411-33.

9. Shahar E, Whitney C, Redline S, Lee E, Newman A, Nieto F, O'Connor G, Boland L, Schwartz J, and Samet J, Sleep disordered breathing and cardiovascular diseases: cross sectional results of the sleep heart health study. *Am J Respir Crit Care Med*, 2001. **163**: 19-25.

10. Peppard P, Young T, Palta M, and Skatrud J, Prospective study of the association between sleep-disordered breathing and hypertension. *NEJM*, 2000. **342**(19): 1378-84.

11. Gander P, Marshall N, Harris R, and Reid P, The Epworth Sleepiness Scale: Influence of age, ethnicity, and socio-economic deprivation. *Sleep*, 2005. **28**(2): 249-53.

12. Wright J, Johns R, Watt I, Melville A, and Sheldon T, Health effects of obstructive sleep apnoea and the effectiveness of continuous positive airway pressure: a systematic review of the research evidence. *BMJ*, 1997. **314**: 851-60.

13. Vgontzas AN, Bixler EO, and Chrousos GP, Sleep apnea is a manifestation of the metabolic syndrome. *Sleep Medicine Reviews*, 2005. **9**(3): 211-24.

14. American Academy of Sleep Medicine Taskforce, Sleep-related breathing disorders in Adults: recommendations for syndrome definition and measurement techniques in clinical research. *Sleep*, 1999. **22**(5): 667-89.

15. Veasey S, Panckeri K, Hoffman E, Pack A, and Hendricks J, The effects of serotonin antagonists in an animal model of sleep disordered breathing. *Am J Respir Crit Care Med*, 1996. **153**: 776-86.

16. Hendricks J and Pack A, *Animal models of sleep apnea*, in *Sleep and Breathing*, N Saunders and C Sullivan, Editors. 1994, Marcel Dekker: New York. p. 847-66.

17. Fogel RB, Malhotra A, and White DP, Sleep 2: Pathophysiology of obstructive sleep apnoea/hypopnoea syndrome. *Thorax*, 2004. **59**(2): 159-63.

18. Jordan A and McEvoy R, Gender differences in sleep apnea: epidemiology, clinical presentation and pathogenic mechanisms. *Sleep Med Rev*, 2003. **7**(5): 377-89.

19. Sforza E, Bacon W, Weiss T, Thibault A, Petiau C, and Krieger J, Upper Airway Collapsibility and Cephalometric Variables in Patients with Obstructive Sleep Apnea. *Am J Respir Crit Care Med*, 2000. **161**(2): 347-52.

20. King ED, O'Donnell CP, Smith PL, and Schwartz AR, A Model of Obstructive Sleep Apnea in Normal Humans . Role of the Upper Airway. *Am J Respir Crit Care Med*, 2000. **161**(6): 1979-84.

21. Fogel R, Amlhotra A, Pillar G, Edwards J, Beauregard J, Shea SA, and White D, Genioglossal activation in patients with obstructive sleep apnea versus controls subjects: mechanisms of muscle control. *Am J Respir Crit Care Med*, 2001. **164**(2025-2030.).

22. Mathur R and Douglas N, Family studies in patients with the sleep apnea-hypopnea syndrome. *Ann Int Med*, 1995. **122**: 174-78.

23. Redline S, Leitner J, Arnold J, Tishler P, and Altose M, Ventilatory-control abnormalities in familial sleep apnea. *Am J Respir Crit Care Med*, 1997. **156**: 155-60.

24. Walshe P, Smith D, Coakeley D, Dunne B, and Timon C, Sleep apnoea of unusual origin. *J Laryngol Otol*, 2002. **116**(2): 138-39.

25. Strohl K and Redline S, Recognition of obstructive sleep apnea. *Am J Respir Crit Care Med*, 1996. **154**: 274-89.

26. Kimoff R, Sforza E, Champagne V, Ofiara L, and Gendron D, Upper airway sensation in snoring and sleep apnea. *Am J Respir Crit Care Med*, 2001. **164**: 2013-14.

27. Gora J, Trinder J, Pierce R, and Colrain I, Evidence of a sleep-specific blunted cortical response to inspiratory occlusions in mild obstructive sleep apnea syndrome. *Am J Respir Crit Care Med*, 2002. **166**: 1225-34.

28. Rubinstein I, Kimoff RJ, Champagne V, and Svanborg E, Upper airway inflammation in obstructive sleep apnea. *Am. J. Respir. Crit. Care Med.*, 2002. **165**(7): 1023-24.

29. Kuna S and Remmers J, *Anatomy and physiology of upper airway obstruction.*, in *Principles and Practice of Sleep Medicine (3rd ed.)*, MH Kryger, T Roth, and WC Dement, Editors. 2000, WB Saunders: Philadelphia.

30. Strohl K, Con: Sleep apnea is an anatomic disorder (Pro/Con Editorial). *Am J Respir Crit Care Med*, 2003. **168**: 270-73.

31. Schwab R, Pro: Sleep apnea is an anatomic disorder (Pro/Con Editorial). *Am J Respir Crit Care Med*, 2003. **168**: 270-73.

32. Young T, Shahar E, Nieto FJ, Redline S, Newman AB, Gottlieb DJ, Walsleben J, Enright P, and Samet J, Predictors of Sleep-Disordered Breathing in Community-Dwelling Adults: The Sleep Heart Health Study. *Arch Int Med*, 2002. **162**: 893-900.

33. Mihaere K, *Obstructive sleep apnoea in Aotearoa/ New Zealand: An objective and questionnaire-based approach to population prevalence estimation and clinical screening*, in *Sleep Wake Research Centre*. 2004, Massey University: Wellington, New Zealand.

34. Coltman R, Taylor R, Whyte K, and Harkness M, Craniofacial form and obstructive sleep apnoea in Polynesian and Caucasian men. *Sleep*, 2000. **23**: 943-50.

35. Duran J, Esnaola S, Rubio R, and Iztueta A, Obstructive Sleep Apnea-Hypopnea and Related Clinical Features in a Population-based Sample of Subjects Aged 30 to 70 Yr. *Am J Respir Crit Care Med*, 2001. **163**(3): 685-89.

36. Bixler E, Vgontzas A, Lin H, Ten Have T, Rein J, Vela-Bueno A, and Kales A, Prevalance of sleep disordered breathing in woman: effects of gender. *Am J Respir Crit Care Med*, 2001. **163**: 608-13.

37. Redline S, Kump K, Tishler P, Browner I, and Ferrette V, Gender differences in sleep disordered breathing in a community-based sample. *Am J Respir Crit Care Med*, 1994. **149**(3): 722-26.

38. Marshall NS, Neill AM, Campbell AJ, and Sheppard DS, Randomised controlled crossover trial of humidified continuous positive airway pressure in mild obstructive sleep apnoea. *Thorax*, 2005. **60**(5): 427-32.

39. Neill A, Wai H, Bannan S, Beasley C, Weatherall M, and Campbell A, Humidified nasal continuous positive airway pressure in obstructive sleep apnoea. *Eur Respir J*, 2003. **22**(2): 258-62.

40. Skinner MA, Kingshott RN, Jones DR, and Taylor DR, Lack of Efficacy for a Cervicomandibular Support Collar in the Management of Obstructive Sleep Apnea. *Chest*, 2004. **125**(1): 118-26.

41. Schwab R, Sex differences and sleep apnoea (editorial). *Thorax*, 1999. **54**: 284-85.

42. Whittle A, Marshall I, Mortimore I, Wraith PK, Sellar R, and Douglas N, Neck soft tissue and fat distribution: comparison between normal men and woman by magnetic resonance imaging. *Thorax*, 1999. **54**(323-328.).

43. Malhotra A, Huang Y, Fogel R, Pillar G, Edwards J, Kikinis R, Loring S, and White D, The male predisposition to pharangeal collapse: The importance of airway length. *Am J Respir Crit Care Med*, 2002. **166**: 1388-95.

44. Bixler E, Vgontzas A, Ten Have T, Tyson K, and Kales A, Effects of age on sleep apnea in men. *Am J Respir Crit Care Med*, 1998. **157**: 144-48.

45. Haas DC, Foster GL, Nieto FJ, Redline S, Resnick HE, Robbins JA, Young T, and Pickering TG, Age-Dependent Associations Between Sleep-Disordered Breathing and Hypertension: Importance of Discriminating Between Systolic/Diastolic Hypertension

and Isolated Systolic Hypertension in the Sleep Heart Health Study. *Circulation*, 2005. **111**(5): 614-21.

46. Paine S-J, Gander P, Harris R, and Reid P, Who reports insomnia? Relationships with age, sex, ethnicity, and socioeconomic deprivation. *Sleep*, 2004. **27**(6): 1163-69.

47. Harris R, *Obstructive sleep apnoea syndrome: symptoms and risk factors among Maori and non-Maori adults in Aotearoa*, in *Department of Public Health*. 2003, University of Otago at the Wellington School of Medicine and health Sciences: Wellington.

48. Statistics New Zealand, *2001 Census of population and dwellings: Maori*. 2002, Statistics New Zealand: Wellington.

49. Ip MSM, Lam B, Lauder IJ, Tsang KWT, Chung K-f, Mok Y-w, and Lam W-k, A Community Study of Sleep-Disordered Breathing in Middle-aged Chinese Men in Hong Kong. *Chest*, 2001. **119**(1): 62-69.

50. Ip MSM, Lam B, Tang LCH, Lauder IJ, Ip TY, and Lam WK, A Community Study of Sleep-Disordered Breathing in Middle-Aged Chinese Women in Hong Kong: Prevalence and Gender Differences. *Chest*, 2004. **125**(1): 127-34.

51. Kim J, In K, Kim J, You S, Kang K, Shim J, Lee S, Lee J, Lee S, Park C, and Shin C, Prevalence of Sleep-disordered Breathing in Middle-aged Korean Men and Women. *Am. J. Respir. Crit. Care Med.*, 2004. **170**(10): 1108-13.

52. Peppard PE, Young T, Palta M, Dempsey J, and Skatrud J, Longitudinal study of moderate weight change and sleep disordered breathing. *JAMA*, 2000. **284**(23): 3015-21.

53. Gottlieb D, Whitney C, Bonekat W, Iber C, James G, Lebowitz M, Nieto F, and Rosenberg C, Relation of sleepiness to respiratory disturbance index: the sleep heart health study. *Am J Respir Crit Care Med*, 1999. **159**: 502-07.

54. Whitney C, Enright P, Newman A, Bonekat W, Foley D, and Quan S, Correlates of daytime sleepiness in 4578 elderly persons: the cardiovascular health study. *Sleep*, 1998. **21**(1): 27-36.

55. Nye E, Paulin J, and Russell D, Blood pressure in a random sample of the New Zealand population. *NZ Med J*, 1992. **105**(926): 1-3.

56. Bullen C, Tipene-Leach D, Vander Hoorn S, Jackson R, Norton R, and MacMahon S, Ethnic differences in blood pressure: findings from the Fletcher Challenge- Auckland University Heart and Health study. *NZ Med J*, 1996. **109**(1032): 395-97.

57. Slight P, *Essential hypertension*, in *Oxford textbook of medicine (2nd ed.)*, D Weatherall, J Ledingham, and D Warrell, Editors. 1987, Oxford University Press: Oxford. p. 13.362-13.82.

58. Hla K, Skatrud J, Finn L, Palta M, and Young T, The effect of correction of sleep disordered breathing on BP in untreated hypertension. *Chest*, 2002. **122**(4): 1125-32.

59. Young T and Javaheri S, *Systemic and Pulmonary Hypertension in Obstructive Sleep Apnea*, in *Principles and Practice of Sleep Medicine (4th ed.)*, MH Kryger, T Roth, and WC Dement, Editors. 2005, Elsevier Saunders: Philadelphia. p. 1192-202.

60. Shepard JJ, *Cardiorespiratory changes in obstructive sleep apnea*, in *Principles and practice of sleep medicine (2nd ed.)*, M Kryger, T Roth, and W Dement, Editors. 1994, WB Saunders: Philadelphia.

61. Tilkian A, Guilleminault C, and Schroeder J, Hemodynamics in sleep-induced apnea. Studies during wakefulness and sleep. *Ann Int Med*, 1976. **85**(714-719.).

62. Przybyowski T, Bangash M-F, Reichmuth K, Morgan BJ, Skatrud JB, and Dempsey JA, Mechanisms of the cerebrovascular response to apnoea in humans. *J Physiology*, 2003. **548**(1): 323-32.

63. Okabe S, Hida W, Kikuchi Y, Taguchi O, Ogawa H, Mizusawa A, Miki H, and Shirato K, Role of hypoxia on increased blood pressure in patients with obstructive sleep apnoea. *Thorax*, 1995. **50**(1): 28-34.

64. Morgan B, Dempsey J, Pegelow D, Jacques A, Finn L, Palta M, Skatrud J, and Young T, Blood pressure pertubations caused by sub-clinical sleep disordered breathing. *Sleep*, 1998. **21**: 737-46.

65. Stradling J and Davies R, Sleep apnea and hypertension- what a mess! *Sleep*, 1997. **20**: 789-93.

66. Wright J and Sheldon T, Sleep apnoea and its impact on public health. *Thorax*, 1998. **53**(5): 410-13.

67. Brooks D, Horner R, Kimoff R, and al. e, Obstructive sleep apnea as a cause of systemic hypertension: Evidence from a canine model. *J Clin Invest*, 1997. **99**: 106-9.

68. Morgan BJ, Crabtree DC, Puleo DS, Badr MS, Toiber F, and Skatrud JB, Neurocirculatory consequences of abrupt change in sleep state in humans. *J Appl Physiol*, 1996. **80**(5): 1627-36.

69. Brooks D, Horner R, Kozar L, and al. e, Effect of obstructive sleep apnea versus sleep fragmentation on responses to airway occlusion. *Am J Respir Crit Care Med*, 1997. **155**: 1609-17.

70. Fletcher E, Lesske J, Qian W, Miller C, and Unger T, Repetitive episodic hypoxia causes diurnal elevation of blood pressure in rats. *Hypertension*, 1992. **19**: 555-61.

71. Bao G, Metreveli N, and Fletcher E, Acute and chronic blood pressure response to recurrent acoustic arousal in rats. *Am J Hypertens.*, 1999. **12**: 504-10.

72. Leung R and Bradley T, Sleep apnea and cardiovascular disease. *Am J Respir Crit Care Med*, 2001. **164**: 2147-65.

73. Bixler E, Vgontzas A, Hung-Mo L, Ten Have T, Leiby B, Velo-Bueno A, and Kales A, Association of hypertension and sleep disordered breathing. *Arch Int Med*, 2000. **160**: 2289-95.

74. Young T, Peppard PE, Palta M, Hla M, Finn L, Morgan B, and Skatrud J, Population based study of sleep disordered breathing as a risk factor for hypertension. *Arch Int Med*, 1997. **157**: 1746-52.

75. Young T and Peppard P, Sleep-disordered breathing and cardiovascular disease: epidemiologic evidence for a relationship. *Sleep*, 2000. **23**(Supp.4): S122-S26.

76. Nieto FJ, Young TB, Lind BK, Shahar E, Samet JM, Redline S, D'Agostino RB, Newman AB, Lebowitz MD, and Pickering TG, Association of sleep-disordered breathing, sleep apnea, and hypertension in a large community-based study. *JAMA*, 2000. **283**(14): 1829-36.

77. Lavie P, Herer P, Peled R, Berger I, Yoffe N, Zomer J, and al. e, Mortality in sleep apnea patients- a multivariate analysis of risk factors. *Sleep*, 1995. **18**(149-157.).

78. Becker H, Jerrentrup A, Ploch T, Grote L, Penzel T, Sullivan C, and Peter J, Effect of nasal continuous positive airway pressure treatment on blood pressure in patients with obstructive sleep apnea. *Circulation*, 2003. **107**: 68-73.

79. Pepperell J, Ramdassingh-Dow S, Crosthwaite N, Mullins R, Jenkinson C, Stradling J, and Davies RJO. Ambulatory blood pressure after therapeutic and sub-therapeutic nasal continuous positive airway pressure for obstructive sleep apnoea: a randomised parallel trial. *Lancet*, 2001. **359**: 204-10.

80. Faccenda J, Mackay T, Boon N, and Douglas N, Randomized placebo-controlled trial of continuous positive airway pressure on blood pressure in the sleep apnea-hypopnea syndrome. *Am J Respir Crit Care Med*, 2001. **163**: 344-48.

81. Engleman HM, Gough K, Martin S, Kingshott RN, Padfield P, and Douglas N, Ambulatory blood pressure on and off continuous positive airway pressure therpay for the apnea/hypopnea syndrome: effects in "non-dippers". *Sleep*, 1996. **19**(5): 378-81.

82. Dimsdale J, Loredo J, and Profant J, Effect of continuous positive airway pressure on blood pressure: a placebo trial. *Hypertension*, 2000. **35**(1): 144-52.

83. Barbe F, Mayoralas L, Duran J, Masa J, Maimo A, Montserrat J, Monasterio C, Bosch M, Ladaria A, Rubio M, Rubio R, Medinas M, Hernandez L, Vidal S, Douglas N, and Agusti A, Treatment with continuous positive airway pressure is not effective in patients with sleep apnea but no daytime sleepiness. *Ann Int Med*, 2001. **134**(11): 1015-67.

84. Redline S, Kapur V, Sanders M, Quan S, Gottlieb D, Rapoport D, Bonekat W, Smith P, Kiley J, and Iber C, Effects of Varying Approaches for Identifying Respiratory Disturbances on Sleep Apnea Assessment. *Am. J. Respir. Crit. Care Med.*, 2000. **161**(2): 369-74.

85. Lavie P and Hoffstein V, Sleep apnea syndrome: a possible contributing factor to resistant hypertension. *Sleep*, 2001. **24**(6): 721-25.

86. Tasali E and Van Couter E, Sleep disordered breathing and the current epidemic of obesity: Consequence or contributing factor? (editorial). *Am J Respir Crit Care Med*, 2002. **165**: 562-63.

87. Ip M, Lam B, Ng M, Lam W, Tsang K, and Lam K, Obstructive sleep apnea is independently associated with insulin resistance. *Am J Respir Crit Care Med*, 2002. **165**: 670-76.

88. Punjabi N, Sorkin J, Katzel L, Goldberg A, Schwartz A, and Smith P, Sleep disordered breathing and insulin resistance in middle-aged and overweight men. *Am J Respir Crit Care Med*, 2002. **165**: 677-82.

89. Gotsopoulos H, Kelly J, and Cistulli P, Oral appliance therapy reduces blood pressure in obstructive sleep apnea: a randomised controlled trial. *Sleep*, 2004. **27**(5): 934-41.

90. Barnes M, McEvoy RD, Banks S, Tarquinio N, Murray CG, Vowles N, and Pierce RJ, Efficacy of Positive Airway Pressure and Oral Appliance in Mild to Moderate Obstructive Sleep Apnea. *Am. J. Respir. Crit. Care Med.*, 2004. **170**(6): 656-64.

91. Peker Y, Hedner J, Norum J, Kraiczi H, and Carlson J, Increased incidence of cardiovascular disease in middle-aged men with obstructive sleep apnea: a 7 year follow-up. *Am J Respir Crit Care Med*, 2002. **166**: 159-65.

92. Silvestrini M, Rizzato B, Placidi F, Baruffaldi R, Bianconi A, and Diomedi M, Carotid artery wall thickness in patients with obstructive sleep apnea syndrome. *Stroke*, 2002. **33**(1782-1785.).

93. Amatoury J, *Snoring Energy Transmission to the Carotid Artery - An Animal Model*, in *Graduate School of Biomedical Engineering*. 2004, University of New South Wales: Sydney.

94. Phillips B, Sleep-disordered breathing and cardiovascular disease. *Sleep Medicine Reviews*, 2005. **9**(2): 131-40.

95. Kanagala R, Murali NS, Friedman PA, Ammash NM, Gersh BJ, Ballman KV, M. Shamsuzzaman AS, and Somers VK, Obstructive Sleep Apnea and the Recurrence of Atrial Fibrillation. *Circulation*, 2003. **107**(20): 2589-94.

96. Dyken M, Somers V, Yamada T, Ren Z-Y, and Zimmerman B, Investigating the relationship between stroke and obstructive sleep apnea. *Stroke*, 1996. **27**: 401-07.

97. Newman A, Nieto F, Guidry U, Lind B, Redline S, Shahar E, Pickering T, and Quan S, Relation of sleep disordered breathing to cardiovascular disease risk factors: the sleep heart health study. *Am J Epidemiol*, 2001. **154**(1): 50-59.

98. Wilcox I, McNamara SG, Collins FL, Grunstein RR, and Sullivan CE, "Syndrome Z": the interaction of sleep apnoea, vascular risk factors and heart disease. *Thorax*, 1998. **53**(90003): 25S-28.

99. Young T. *Epidemiological evidence for a role for OSA in cardiovascular disease.* in *Annual Scientific Meeting of the Australasian Sleep Association.* 2004. Sydney, Australia.

100. Doherty LS, Kiely JL, Swan V, and McNicholas WT, Long-term Effects of Nasal Continuous Positive Airway Pressure Therapy on Cardiovascular Outcomes in Sleep Apnea Syndrome. *Chest*, 2005. **127**(6): 2076-84.

101. Marin J, Carrizo S, Vicente E, and Agusti A, Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. *Lancet*, 2005. **365**: 1046-53.

102. Belenky G, Wesensten N, Thorne D, Thomas M, Sing H, Redmond D, Russo M, and Balkin T, Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *J Sleep Res*, 2003. **12**(1): 1-12.

103. Van Dongen H, Maislin G, Mullington J, and Dinges DF, The Cumulative Cost of Additional Wakefulness: Dose-Response Effects on Neurobehavioral Functions and Sleep Physiology From Chronic Sleep Restriction and Total Sleep Deprivation. *Sleep*, 2002. **26**(2): 117-26.

104. Stepanski E, The effect of sleep fragmentation on daytime function. *Sleep*, 2002. **25**(3): 268-76.

105. Tietzel AJ and Lack LC, The recuperative value of brief and ultra-brief naps on alertness and cognitive performance. *J Sleep Res*, 2002. **11**(3): 213-18.

106. Wesensten NJ, Balkin TJ, and Belenky G, Does sleep fragmentation impact recuperation?A review and reanalysis. *J Sleep Res*, 1999. **8**(4): 237-45.

107. Kim H, Young T, Matthews C, Weber S, Woodard A, and Palta M, Sleep-disordered breathing and neuropsychological deficits: a population based study. *Am J Respir Crit Care Med*, 1997. **156**(6): 1813-19.

108. Adams N, Strauss ME, Schluchter M, and Redline S, Relation of measures of sleep disordered breathing to neuropsychological functioning. *Am J Respir Crit Care Med*, 2001. **163**: 1626-31.

109. Kingshott RN, Vennelle M, Hoy CJ, Engleman HM, Deary IJ, and Douglas NJ, Predictors of improvements in daytime function outcomes with CPAP therapy. *Am J Respir Crit Care Med*, 2000. **161**: 866-71.

110. Kingshott RN, Engleman HM, Deary IJ, and Douglas NJ, Does arousal frequency predict daytime function? *Eur Respir J*, 1998. **12**: 1264-70.

111. National Health and Medical Research Council of Australia. *Effectiveness of nasal continuous positive airway pressure (nCPAP) in obstructive sleep apnoea in adults.* 2000, NHMRC: Canberra. 129.

112. Kazis L, Anderson J, and Meenan R, Effect sizes for interpreting changes in health status. *Med Care*, 1989. **27**(3 suppl): S178-S89.

113. Engleman HM, Kingshott RN, Martin SE, and Douglas NJ, Cognitive function in the sleep apnea/hypopnea syndrome (SAHS). *Sleep*, 2000. **23**(supp 4): S102-S08.

114. Bedard M, Montplaisir J, Richer F, Rouleau I, and Malo J, Obstructive sleep apnea syndrome: pathogenesis of neuropsychological deficits. *Journal of Clinical & Experimental Neuropsychology*, 1991. **13**(6): 950-64.

115. Kamba M, Suto Y, Ohta Y, Inoue Y, and Matsuda E, Cerebral metabolism in sleep apnea: evaluation by magnetic resonance spectroscopy. *American journal of respiratory and critical care medicine*, 1997. **156**(296-298.).

116. Bartlett DJ, Rae C, Thompson CH, Byth K, Joffe DA, Enright T, and Grunstein RR, Hippocampal area metabolites relate to severity and cognitive function in obstructive sleep apnea. *Sleep Medicine*, 2004. **5**(6): 593-96.

117. Macey P, Henderson L, Macey K, Alger J, Frysinger R, Woo M, Harper R, Yan-Go F, and Harper R, Brain morphology associated with obstructive sleep apnea. *Am J Respir Crit Care Med*, 2002. **166**: 1382-87.

118. Gozal D, The brain is sleep disordered breathing: Is it the chicken or the egg? (editorial). *Am J Respir Crit Care Med*, 2002. **166**: 1305-09.

119. Young T, Blustein J, Finn L, and Palta M, Sleep-disordered breathing and motor vehicle accidents in a population based sample of employed adults. *Sleep*, 1997. **20**: 608-13.

120. Johns MW, Sleep propensity varies with behaviour and the situation in which it is measured: The concept of somnificity. *J Sleep Res*, 2002. **11**: 61-67.

121. Marshall N, Bolger W, and Gander P, Abnormal sleep duration and motor vehicle crash risk (research letter). *J Sleep Res*, 2004. **13**(2): 177-78.

122. Gander P, Marshall N, Harris R, and Reid P, Sleep, sleepiness and motor vehicle accidents: a national survey. *Aust NZ J Pub Heal*, 2005. **29**: 16-21.

123. Marshall N, Gander P, and Neill A, Obstructive sleep apnoea and motor vehicle accident risk: a perspective. *NZ Med J*, 2003. **116**(1176): 482-89.

124. Findley L, Fabrizio M, Thommi G, and Suratt P, Severity of sleep apnoea and automobile crashes (letter). *NEJM*, 1989. **320**: 868-69.

125. George CFP, Reduction in motor vehicle collisions following treatment of sleep apnoea with nasal CPAP. *Thorax*, 2001. **56**: 508-12.

126. Findley L, Smith C, Hooper J, Dineen M, and Suratt P, Treatment with nasal CPAP decreases automobile accidents in patients with sleep apnea. *Sleep*, 2000. **161**: 857-59.

127. Lindberg E, Carter N, Gislason T, and Janson C, Role of snoring and daytime sleepiness in occupational accidents. *Am J Respir Crit Care Med*, 2001. **164**: 2031-35.

128. Robertson C, Herbison P, and Harkness M, Dental and occlusal changes during mandibular advancement splint therapy in sleep disordered patients. *Eur J Orth*, 2003. **25**(4): 371-76.

129. Fergusson K and Lowe A, *Oral appliances for sleep disordered breathing*, in *Principles and Practice of Sleep Medicine (4th ed.)*, MH Kryger, T Roth, and WC Dement, Editors. 2005, Elsevier Saunders: Philadelphia PA. p. 1098-108.

130. Lim J, Lasserson T, Fleetham J, and Wright J, Oral appliances for obstructive sleep apnoea (Cochrane Review). *Cochrane Database Syst Rev*, 2004. **4**: CD004435.

131. Clark G, Blumenfeld I, Yoffe N, Peled E, and Lavie P, A crossover study comparing the efficacy of continuous positive airway pressure with anterior mandibular positioning devices on patients with obstructive sleep apnea. *Chest*, 1996. **109**: 1477-83.

132. Fergusson K, Ono T, Lowe A, Keenan S, and Fleetham J, A randomised crossover study of an oral appliance vs nasal-continuous positive airway pressure in the treatment of mild-moderate obstructive sleep apnea. *Chest*, 1996. **109**: 1269-75.

133. Riley RW, Powell NB, Kasey KL, and Guilleminault C, *Surgical therapy for obstructive sleep apnea-hypopnea syndrome*, in *Principles and practice of sleep medicine (3rd ed.)*, MH Kryger, T Roth, and WC Dement, Editors. 2000, W.B. Saunders: Philadelphia. p. 913-28.

134. Haapaniemi J, Laurikainen E, Halme P, and Antila J, Long-term results of tracheostomy for severe obstructive sleep apnea syndrome. *ORL: Journal of Oto-Rhino-Laryngology & its Related Specialties*, 2001. **63**(3): 131-6.

135. Powell N, Riley R, and Guilleminault C, *Surgical management of sleep-disordered breathing*, in *Principles and Practice of Sleep Medicine (4th ed)*, MH Kryger, T Roth, and WC Dement, Editors. 2005, Elsevier Saunders: Philadelphia PA.

136. Bridgman S, Dunn K, and Ducharme F, Surgery for obstructive sleep apnoea (Cochrane Review). *Cochrane Database of Systematic Reviews*, 2003. **3**.

137. Valencia-Flores M, Orea A, Herrera M, Santiago V, Rebollar V, Castano V, Oseguera J, Pedroza J, Sumano J, Resendiz M, and Garcia-Ramos G, Effect of bariatric surgery on obstructive sleep apnea and hypopnea syndrome, electrocardiogram, and pulmonary arterial pressure. *Obesity Surgery*, 2004. **14**(6): 755-62.

138. Sanders MH, *Medical therapy for obstructive sleep apnea-hypopnea syndrome*, in *Principles and practice of sleep medicine (3rd ed.)*, MH Kryger, T Roth, and WC Dement, Editors. 2000, W.B. Saunders: Philadelphia. p. 879-93.

139. Shneerson J and Wright J, *Lifestyle modification for obstructive sleep apnoea (Cochrane Review)*. 2000, The Cochrane Library: Oxford: Update Software.

140. Sullivan CE, Issa FG, Bethon-Jones M, and Eves L, Reversal of obstructive sleep apnoea by continuous positive airway pressure applied through the nares. *Lancet*, 1981. **1**: 862-65.

141. Pack A and Young T, Superficial analysis ignores evidence on efficacy of treatment (letter). *BMJ*, 1997. **315**: 367.

142. Engleman H, Martin S, Deary I, and Douglas N, Effect of nasal continuous positive airway pressure treatment on daytime function in sleep apnoea/hypopnoea syndrome. *Lancet*, 1994. **343**: 572-75.

143. Shneerson J, Smith I, Pack A, Young T, Stradling J, Davies R, Gibson G, Prowse K, Semple S, London D, Engleman H, Martin S, Deary I, Douglas N, Wright J, and Sheldon T, Obstructive sleep apnoea (letters). *BMJ*, 1997. **315**(7104): 367 - 67.

144. Engleman H, Martin S, Deary I, and Douglas N, Some criticisms of studies are unfounded (letter). *BMJ*, 1997. **315**: 367.

145. Stradling J and Davies R, Evidence for efficacy of continuous positive airways pressure is compelling (letter). *BMJ*, 1997. **315**: 367.

146. Stradling JR and Davies RJO, The unacceptable face of evidence-based medicine. *J Eval Clin Pract*, 1997. **3**(2): 99-103.

147. Stradling JR, Sleep apnoea and the misuse of evidence based medicine. *Lancet*, 1997. **349**: 201-02.

148. Australian Health and Technology Advisory Committee, *The effectiveness and cost-effectiveness of nasal continuous positive airway pressure in the treatment of obstructive sleep apnea in adults*. 1996, Australian Health Technology Advisory Committee: Canberra.

149. Patel S, White D, Malhotra A, Stanchina M, and Ayas N, Continuous positive airway pressure therapy for treating sleepiness in a diverse population with obstructive sleep apnea: results of a meta-analysis. *Arch Int Med*, 2003. **163**: 565-71.

150. Jenkinson C, Davies RJO, Mullins R, and Stradling JR, Comparison of therapeutic and subtherapeutic nasal continuous positive airway pressure for obstructive sleep apnoea: a randomised prospective parallel trial. *Lancet*, 1999. **353**: 2100-05.

151. Engleman H, Martin S, Kingshott R, Mackay T, Deary I, and Douglas N, Randomised placebo controlled trial of daytime function after continuous positive airway pressure (CPAP) therapy for the sleep apnoea/hypopnoea syndrome. *Thorax*, 1998. **53**: 341-45.

152. Sangal R, Thomas L, and Mitler MM, Maintenance of wakefulness test and multiple sleep latency test: measurement of different abilities in patients with sleep disorders. *Clin Neurophysiol*, 1992. **110**(1999): 2131-35.

153. Sangal R, Thomas L, and Mitler MM, Disorders of excessive sleepiness: treatment improves ability to stay awake but does not reduce sleepiness. *Chest*, 1992. **102**(3): 699-703.

154. Dinges DF and Powell NB, Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behav Res Methods Instrum Comput*, 1985. **17**(6): 652-55.

155. Henke KG, Grady JJ, and Kuna ST, Effect of nasal continuous positive airway pressure on neuropsychological function in sleep apnea-hypopnea syndrome: a randomized placebo-controlled trial. *Am J Respir Crit Care Med*, 2001. **163**: 911-17.

156. Hack M, Davies RJO, Mullins R, Choi SJ, Ramdassingh-Dow S, Jenkinson C, and Stradling JR, Randomised prospective trial of therapeutic vs. sub-therapeutic nasal continuous airway pressure on simulated steering performance in patients with obstructive sleep apnoea. *Thorax*, 2000. **55**: 224-31.

157. Bardwell W, Ancoli-Israel S, Berry C, and Dimsdale J, Neuropsychological effects of one week continuous positive airway pressure treatment in patients with obstructive sleep apnea: a placebo controlled study. *Psychsom Med*, 2001. **63**: 579-84.

158. Priest B, Brichard C, Aubert G, Liistro G, and Rodenstein DO, Microsleep during a simplified maintenance of wakefulness test: a validation of the OSLER test. *Am J Respir Crit Care Med*, 2001. **163**: 1619-25.

159. Turkington PM, Sircar M, Allgar V, and Elliot MW, Relationship between obstructive sleep apnoea, driving simulator performance, and risk of road traffic accidents. *Thorax*, 2001. **56**(10): 800-05.

160. Banks S, Catchside P, Lack L, Grunstein R, and McEvoy R, Low levels of alcohol impair driving simulator performance and reduce perception of crash risk in partially sleep deprived subjects. *Sleep*, 2004. **27**(6): 1063-7.

161. Ballester E, Badia J, Hernandez L, Carasco E, de Pablo J, Fornas C, Rodriquez-Roisin R, and Montserrat J, Evidence of the effectiveness of continuous positive airway pressure in the treatment of sleep apnea/hypopnea syndrome. *Am J Respir Crit Care Med*, 1999. **159**: 495-501.

162. Yu B-H, Ancoli-Israel S, and Dimsdale JE, Effect of CPAP treatment on mood states in patients with sleep apnea. *J Psych Res*, 1999. **33**: 427-32.

163. Profant J, Ancoli-Israel S, and Dimsdale JE, A randomized, controlled trial of 1 week of continuous positive airway pressure treatment on quality of life. *Heart Lung*, 2003. **32**(1): 52-58.

164. Montserrat JM, Ferrer M, Hernandez L, Farre R, Vilagut G, Navajas D, Badia JR, Carrasco E, de Pablo J, and Ballester E, Effectiveness of CPAP treatment in daytime function in sleep apnea syndrome: a randomized controlled study with an optimized placebo. *Am J Respir Crit Care Med*, 2001. **164**: 608-13.

165. Schultz K and Grimes D, Generation of allocation sequences in randomised trials: chance, not choice. *Lancet*, 2002. **359**: 515-19.

166. Schultz K and Grimes D, Allocation concealment in randomised trials: defending against deciphering. *Lancet*, 2002. **359**(9306): 614-18.

167. Schultz K, Altman D, and Moher D, Allocation concealment in clinical trials (letter). *JAMA*, 2002. **288**(9).

168. Karlawish JHT and Pack AI, Addressing the ethical problems of randomized and placebo controlled trials of CPAP. *Am J Respir Crit Care Med*, 2001. **163**: 809-10.

169. Farre R, Hernandez L, Montserrat J, Rotger M, Ballester E, and Navajas D, Sham continuous positive airway pressure for placebo-controlled studies in sleep apnoea. *Lancet*, 1999. **353**: 1154.

170. Douglas N, Engleman H, Faccenda J, and McArdle N, The science of designing ethical CPAP trials (letter). *Am J Respir Crit Care Med*, 2002. **165**: 132-33.

171. Barnes M, Houston D, Worsnop C, Neill A, Mykytyn I, Kay A, Trindler J, Saunders N, McEvoy R, and Pierce R, A randomized controlled trial continuous positive airway pressure in mild obstructive apnea. *Am J Respir Crit Care Med*, 2002. **165**: 773-80.

172. Beck J and Azari E, FDA, Off label use and informed consent: debunking myths and misconceptions. *Food and Drug Law Journal*, 1998. **59**(1): 71-104.

173. Whittington C, Kendall T, Fonagy P, Cottrell D, Cotgrove A, and Boddington E, Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet*, 2004. **363**(9418): 1341-45.

174. Jick H, Kaye JA, and Jick SS, Antidepressants and the Risk of Suicidal Behaviors. *JAMA*, 2004. **292**(3): 338-43.

175. Editor L, Depressing Research. *Lancet*, 2004. **363**(9418): 1335.

176. Loube D, Gay P, Strohl K, Pack A, White D, and Collop N, Consensus statment: indications for positive airway pressure treatment of adult obstructive sleep apnea patients. *Chest*, 1999. **115**(3): 863-66.

177. Laitenen L, Anttalainen U, Pietinalho A, Hamalainen P, and Koskela K, Sleep apnoea: Finnish national guidelines for prevention and treatment 2002-2012. *Resp Med*, 2003. **97**: 337-65.

178. Massie C, Hart R, Peralez K, and Richards G, Effects of humidification on nasal symptoms and compliance in sleep apnea patients using continuous positive airway pressure.[comment]. *Chest*, 1999. **116**: 403-08.

179. Sackett D, Rosenberg W, Gray J, Haynes R, and Richardson W, Evidence based medicine: what it is and what it isn't. *BMJ*, 1996. **312**: 71-72.

180. Patsopoulos NA, Analatos AA, and Ioannidis JPA, Relative Citation Impact of Various Study Designs in the Health Sciences. *JAMA*, 2005. **293**(19): 2362-66.

181. Dawkins R, *A Devil's Chaplain*. 2003, New York: Houghton Mifflin.

182. Ellis J, Mulligan I, Rowe J, and Sackett D, Inpatient general medicine is evidence based. *Lancet*, 1995. **346**(8972): 407-10.

183. Quan S, Howard B, Iber C, Kiley J, Nieto FJ, O'Connor G, Rapoport D, Redline S, Robbins J, Samet J, and Wahl P, The sleep heart health study: design, methods and rationale. *Sleep*, 1997. **20**(12): 1077-85.

184. Engleman H, Kingshott R, Wraith P, Mackay T, Deary I, and Douglas N, Randomised placebo controlled crossover trial of continuous positive airway pressure for mild sleep apnea/hypopnea syndrome. *Am J Respir Crit Care Med*, 1999. **159**: 461-67.

185. Engleman H, Martin S, Deary I, and Douglas N, Effect of CPAP therapy on daytime function in patients with mild sleep apnoea/hypopnoea syndrome. *Thorax*, 1997. **52**: 114-19.

186. Monasterio C, Vidal S, Duran J, Ferrer M, Carmona C, Barbe F, Mayos M, Gonzalez-Mangado N, Juncadella M, Navarro A, Barreira R, Capote F, Mayoralas L, Peces-Barba G, Alonso J, and Montserrat J, Effectiveness of continuous positive airway pressure in mild sleep apnea-hypopnea syndrome. *Am J Respir Crit Care Med*, 2001. **164**: 939-43.

187. Redline S, Adams N, Strauss ME, Roebuck T, Winters M, and Rosenberg C, Improvement of mild sleep disordered breathing with CPAP compared with conservative therapy. *Am J Respir Crit Care Med*, 1998. **157**: 858-65.

188. Van Dongen H, Baynard M, Maislin G, and Dinges D, Systematic interindividual differences in neurobehavioural impairment from sleep loss: evidence of trait-like differential vulnerability. *Sleep*, 2004. **27**(3): 423-33.

189. Lamphere J, Roehrs T, Wittig R, Zorick F, Conway W, and Roth T, Recovery of alertness after CPAP in apnea. *Chest*, 1989. **96**: 1364-67.

190. Iber C, Redline S, Kaplan Gilpin A, Quan S, Zhang L, Gottlieb D, Rapoport D, Resnick H, Sanders M, and Smith P, Polysomnography performed in the unattended home versus the attended laboratory setting- sleep heart health study methodology. *Sleep*, 2004. **27**(3): 536-40.

191. Rechtschaffen A and Kales A, *A Manual of standardized terminology, techniques and scoring system for sleep stages of human subjects.* 1968, Brain information service/ Brain research Institute University of California.: Los Angeles.

192. American Sleep Disorders Association, EEG arousals: scoring rules and examples. *Sleep*, 1992. **15**: 173-84.

193. Johns MW, A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*, 1991. **14**: 540-45.

194. Mitler MM, Gujavarty KS, and Browman CP, Maintenance of wakefulness test: a polysomnographic technique for evaluating treatment efficacy in patients with excessive somnolence. *Electroenceph Clin Neurophys*, 1982. **53**(6): 658-61.

195. Doghramji K, Mitler M, Sangal R, Shapiro C, Taylor S, Walsleben J, Belisle C, Erman M, Hayduk R, Hosn R, O'Malley E, Sangal J, Schutte S, and Youakim J, A normative study of the maintenance of wakefulness test (MWT). *Electroenceph Clin Neurophys*, 1997. **103**: 554-62.

196. Banks S, Barnes M, Tarquinio N, Pierce R, Lack L, and McEvoy R, The maintenance of wakefulness test in normal healthy subjects. *Sleep*, 2004. **27**(4): 799-802.

197. Banks S, Kenner A, Barnes M, Pierce R, and McEvoy R. *Can the maintenance of wakefulness test be shortened? A preliminary investigation.* in *Annual Scientific Meeting of the Australasian Sleep Association.* 2000. Melbourne Australia.

198. Weaver TE, Laizner AM, Evans LK, Maislin G, Chugh DK, Lyon K, Smith PL, Schwartz AR, Redline S, Pack AI, and Dinges DF, An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep*, 1997. **20**(10): 835-43.

199. Stewart A, Hays R, and Ware J, The MOS short-form general health survey: reliability and validity in a patient population. *Med Care*, 1988. **26**(7): 724-35.

200. Ware J and Sherbourne C, The MOS 36-item short-form health survey (SF-36): Conceptual framework and item selection. *Med Care*, 1992. **30**(6): 473-83.

201. Ware J and Gandek B, Overview of the SF-36 health survey and the international quality of life assessment (IQOLA) project. *J Clin Epidemiol*, 1998. **51**(11): 903-12.

202. Ware J, Kosinski M, Gandek B, Aaronson N, Apolone G, Bech P, Brazier J, Bullinger M, Kaasa S, Leplege A, Prieto L, and Sullivan M, The factor structure of the SF-36 health survey in 10 countries: results from the IQOLA project. *J Clin Epidemiol*, 1998. **51**(11): 1159-65.

203. McHorney CA, Ware J, Lu JF, and Sherbourne CD, The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care*, 1994. **32**(1): 40-66.

204. Scott K, Tobias MI, Sarfati D, and SJ H, SF-36 health survey reliability, validity, and norms for New Zealand. *Aust NZ J Pub Heal*, 1999. **23**: 401-06.

205. Zigmond A and Snaith R, The hospital anxiety and depression scale. *Acta Psychiatr Scan*, 1983. **67**: 367-70.

206. Mykletun A, Stordal E, and Dahl A, Hospital anxiety and depression scale: factor structure, item analyses and internal consistency in a large population. *Brit J Psych*, 2001. **179**(6): 540-49.

207. Spinhoven P, Ormel J, Sloekers P, Kempen G, Speckens A, and Van Hemert A, A validation of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psych Med*, 1997. **27**: 363-70.

208. Chugh DK, Weaver TE, and Dinges DF, Psychomotor vigilance performance in sleep apnea patients compared to patients presenting with snoring without apnea. *Sleep*, 1998. **21**(S1): 491.K1.

209. Castelloe J and O'Brien R, *Power and sample size determination for linear models*. Unknown, SAS Institute: Cary, NC, USA.

210. Engleman H, When Does `Mild' Obstructive Sleep Apnea/Hypopnea Syndrome Merit Continuous Positive Airway Pressure Treatment? *Am J Respir Crit Care Med*, 2002. **165**(6): 743-45.

211. Moher D, Cook D, Eastwood S, Olkin I, Rennie D, Stroup D, and group fiQ, Improving the quality of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet*, 1999. **354**: 1896-900.

212. White J, Cates C, and Wright J, *Continuous positive airways pressure for obstructive sleep apnoea (Cochrane Review)*. 2000, The Cochrane Library: Oxford: Update Software.

213. Douglas N, Systematic review of the efficacy of Nasal CPAP. *Thorax*, 1998. **53**: 414-15.

214. Mitler MM, Carskadon MA, and Hirshkowitz M, *Evaluating sleepiness*, in *Principles and practice of sleep medicine (3rd ed.)*, MH Kryger, T Roth, and WC Dement, Editors. 2000, WB Saunders Co.: USA.

215. Bonnet MH and Arand D, Arousal components which differentiate the MWT from the MSLT. *Sleep*, 2001. **24**(4): 441-47.

216. Carskadon M, Dement W, Mitler M, Roth T, Westbrook P, and Keenan S, Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness. *Sleep*, 1986. **9**: 519-24.

217. Jadad A, *Randomised controlled trials*. 1998, London: BMJ Books. 123.

218. Banks S, Barnes M, Tarquinio N, Pierce R, Lack L, and McEvoy R, Factors associated with maintenance of wakefulness test mean sleep latency in patients with mild to moderate obstructive sleep apnoea and normal subjects. *J Sleep Res*, 2004. **13**(2): 71-78.

219. Loredo JS, Ancoli-Israel S, and Dimsdale JE, Effect of continuous positive airway pressure vs. placebo continuous positive airway pressure on sleep quality in obstructive sleep apnea. *Chest*, 1999. **116**: 1545-49.

220. Bao X, Nelesen R, Loredo J, Dimsdale J, and Ziegler M, Blood pressure variability in obstructive sleep apnea: role of sympathetic nervous activity and effect of continuous positive airway pressure. *Blood Pressure Monitoring*, 2002. **7**(6): 301-07.

221. Ziegler M, Mills P, Loredo J, Ancoli-Israel S, and Dimsdale J, Effect of continuous positive airway pressure and placebo treatment on sympathetic nervous activity in patients with obstructive sleep apnea. *Chest*, 2001. **120**(3): 887-93.

222. Jokic R, Klimaszewski A, Crossley M, Sridhar G, and Fitzpatrick M, Positional treatment vs continuous positive airway pressure in patients with positional obstructive sleep apnea syndrome. *Chest*, 1999. **115**(3): 771-81.

223. Pelletier-Fleury N, Meslier N, Gagnadoux F, Person C, Rakotonarahary D, Ouksel H, Fleury B, and Racineaux J-L, Economic arguments for the immediate management of moderate to severe obstructive sleep apnoea syndrome. *Eur Respir J*, 2004. **23**: 53-60.

224. Ip MSM, Tse H-F, Lam B, Tsang KWT, and Lam W-K, Endothelial Function in Obstructive Sleep Apnea and Response to Treatment. *Am. J. Respir. Crit. Care Med.*, 2004. **169**(3): 348-53.

225. Engleman HM and Wild M, Improving CPAP use by patients with the sleep apnoea/hypopnoea syndrome (SAHS). *Sleep Med Rev*, 2003. **7**(1): 81-99.

226. Kribbs N, Pack AI, Kline L, Smith P, Schwartz AR, Schubert N, Redline S, Henry J, Getsy J, and Dinges DF, Objective measurement of patterns of nasal CPAP use by patients with obstructive sleep apnea. *Am J Respir Crit Care Med*, 1993. **147**: 887-95.

227. Wiest G, Fuchs F, Brueckl W, Nusko G, Harsch I, Hahn E, and Ficker J, In vivo efficacy of heated and non-heated humidifiers during nasal continuous positive airway pressure (nCPAP)-therapy for obstructive sleep apnoea. *Respir Med*, 2000. **94**(4): 364-68.

228. Grunstein R, *Continuous positive airway pressure treatment for obstructive sleep apnea-hypopnea syndrome*, in *Principles and practice of sleep medicine (4th ed.)*, M Kryger, T Roth, and W Dement, Editors. 2005, Elsevier Saunders: Philadelphia. p. 1066-80.

229. Berry R, Parish J, and Hartse K, The use of auto-titrating continuous positive airway pressure for treatment of adult obstructive sleep apnea. *Sleep*, 2002. **25**(2): 148-59.

230. Farre R, Montserrat JM, Rigau J, Trepat X, Pinto P, and Navajas D, Response of automatic continuous positive airway pressure devices to different sleep breathing patterns: a bench study. *Am J Respir Crit Care Med*, 2002. **166**: 469-73.

231. Littner M, Hirshkowitz M, Davilla D, Anderson W, Kushida C, Woodson B, Johnson S, and Wise M, Practice parameters for the use of auto-titrating continuous positive airway pressure devices for titrating pressures and treating adult patients with obstructive sleep apnea syndrome: an American Academy of Sleep Medicine Report. *Sleep*, 2002. **25**(2): 143-47.

232. Ayas N, White D, Manson J, Stampfer M, Speizer F, Malhotra A, and Hu F, A prospective study of sleep duration and coronary heart disease in women. *Arch Int Med*, 2003. **163**: 205-09.

233. Noseda A, Kempenaers C, Kerkhofs M, Braun S, Linkowski P, and Jann E, Constant vs Auto-Continuous Positive Airway Pressure in Patients With Sleep Apnea Hypopnea Syndrome and a High Variability in Pressure Requirement. *Chest*, 2004. **126**(1): 31-37.

234. Resta O, Guido P, Picca V, Sabato R, Rizzi M, Scarpelli F, and Sergi M, Prescription of nCPAP and nBIPAP in obstructive sleep apnoea syndrome: Italian experience in 105 subjects. A prospective two centre study. *Resp Med*, 1998. **92**: 820-27.

235. Schafer H, Ewig S, Hasper E, and Luderitz B, Failure of CPAP therapy in obstructive sleep apnoea syndrome: predictive factors and treatment with bilevel-positive airway pressure. *Resp Med*, 1998. **92**: 208-15.

236. Reeves-Hoche M, Hudgel D, Meck R, Witteman R, Ross A, and Zwillich C, Continuous versus bilevel positive airway pressure for obstructive sleep apnea. *Am. J. Respir. Crit. Care Med.*, 1995. **151**(2): 443-49.

237. Gay P, Herold D, and Olson E, A randomized, double-blind clinical trial comparing continuous positive airway pressure with a novel bilevel pressure system for treatment of obstructive sleep apnea syndrome. *Sleep*, 2003. **26**(7): 864-69.

238. Aloia MS, Stanchina M, Arnedt JT, Malhotra A, and Millman RP, Treatment Adherence and Outcomes in Flexible vs Standard Continuous Positive Airway Pressure Therapy. *Chest*, 2005. **127**(6): 2085-93.

239. Sanders M, Costantino J, Strollo P, Studnicki K, and Atwood C, Jr, The Impact Of Split-Night Polysomnography For Diagnosis And Positive Pressure Therapy Titration On Treatment Acceptance And Adherence In Sleep Apnea/Hypopnea. *Sleep*, 2000. **23**(1): 17-24.

240. Rodway G and Sanders M, The efficacy of split-night sleep studies. *Sleep Medicine Reviews*, 2003. **7**(5): 391-401.

241. McArdle N, Grove A, Devereux G, Mackay-Brown L, Mackay T, and Douglas N, Split-night versus full-night studies for sleep apnoea/hypopnoea syndrome. *Eur Respir J*, 2000. **15**(4): 670-75.

242. Elshaug AG, Moss JR, and Southcott AM, Implementation of a split-night protocol to improve efficiency in assessment and treatment of obstructive sleep apnoea. *Intern Med J*, 2005. **35**(4): 251-54.

243. Yamashiro Y and Kryger M, CPAP titration for sleep apnea using a split-night protocol. *Chest*, 1995. **107**(1): 62-66.

244. Nelesen R, Yu H, Ziegler M, Mills P, Clausen J, and Dimsdale JE, Continuous positive airway pressure normalizes cardiac autonomic and hemodynamic responses to a laboratory stressor in apneic patients. *Chest*, 2001. **119**: 1092-101.

245. Espana R and Scammell T, Sleep neurobiology for the clinician. *Sleep*, 2004. **27**(4): 811-20.

246. Loube D and Ball N, Comparison of compliance with proportional positive airway pressure (C-Flex) to continuous positive airway pressure treatment in obstructive sleep apnea. *Sleep*, 2004. **27**: A228.

247. Dorrian J, Rogers N, and Dinges D, *Behavioral Alertness as Measured by Psychomotor Vigilance Performance:*

A Neurocognitive Assay Sensitive to the Effects of Sleep Loss, in *Sleep Deprivation: Clinical Issues, Pharmacology, And Sleep Loss Effects*, C Kushida, Editor. 2004, Marcel Dekker: New York. p. 39-70.

248. Stepnowsky CJ and Moore P, Nasal CPAP treatment for obstructive sleep apnea: developing a new perspective on dosing strategies and compliance. *J Psychsom Res*, 2003. **54**: 599-605.

249. Hers V, Liistro G, Dury M, Collard P, Aubert G, and Rodenstein D, Residual effect of nCPAP applied for part of the night in patients with obstructive sleep apnoea. *Eur Respir J*, 1997. **10**(5): 973-76.

250. McNicholas W, Compliance with nasal CPAP therapy for obstructive sleep apnoea: how much is enough? *Eur Respir J*, 1997. **10**(5): 969-70.

251. Gottlieb DJ, Yao Q, Redline S, Ali T, and Mahowald MW, Does snoring predict sleepiness independently of apnea and hypopnea frequency? *Am J Respir Crit Care Med*, 2000. **162**: 1512-17.

252. Van Dongen H, Doran S, and Dinges DF, Performance during sleep deprivation: evidence for state instability and trait vulnerability. *Sleep*, 2000. **23**(supp 2): A245-A46.

253. Frey DJ, Badia P, and Wright KP, Inter- and intra-individual variability in performance near the circadian nadir during sleep deprivation. *J Sleep Res*, 2004. **13**(4): 305-15.

254. Veasey S, Davis C, Fenik P, Zhan G, Hsu Y-J, Pratico D, and Gow A, Long-term intermittant hypoxia in Mice: protracted hypersomnolence with oxidative injury to sleep-wake brain regions. *Sleep*, 2004. **27**(2): 194-201.

255. Lojander J, Kajaste S, Maasilta P, and Partinen M, Cognitive function and treatment of obstructive sleep apnea syndrome. *J Sleep Res*, 1999. **8**(1): 71-76.

256. Lojander J, Maasilta P, Partinen M, Brander PE, Tapani S, and Lehtonen H, Nasal-CPAP, surgery, and conservative management for treatment of obstructive sleep apnea syndrome. *Chest*, 1996. **110**: 114-19.

257. Marti S, Sampol G, Munoz X, Torres F, Roca A, Lloberes P, Sagales T, Quesada P, and Morell F, Mortality in severe sleep apnoea/hypopnoea syndrome patients: impact of treatment. *Eur Respir J*, 2002. **20**: 1511-18.

258. Beninati W and Sanders M, Optimal continuous positive airway pressure for the treatment of obstructive sleep apnea/hypopnea syndrome. *Sleep Med Rev*, 2001. **5**(1): 7-23.

259. Ayas N, Patel S, Malhotra A, Schulzer M, Malhotra M, Jung D, Fleetham JA, and White D, Auto-Titrating versus standard continuous positive airway pressure for the treatment of obstructive sleep apnea: results of a meta-analysis. *Sleep*, 2004. **27**(2): 249-53.

APPENDIX 1
OUTCOME MEASURES

1.1 EPWORTH SLEEPINESS SCALE

The Epworth Sleepiness Score (ESS) is a simple self administered questionnaire developed in Melbourne, Australia by Murray Johns[3]. It aims to measure chronic subjective sleepiness. The ESS measures self-reported behaviour in 8 given situations with each question scored between 0 (would never doze) and 3 (high chance of dozing). Thus it has a possible range of scores between 0 and 24 (indicating very sleepy). The ESS is quick and easy to administer and is simple to score, which is part of the appeal of this measure and one of the reasons why it is in widespread use. It measures the self-rated likelihood of falling asleep in recent times in some common everyday situations. A higher ESS score indicates a subjectively sleepier person. Scores above 10 are usually regarded as a sign of abnormal sleepiness. Scores of 16 or above are regarded as evidence of extreme sleepiness.

As a pencil and paper device, the ESS is susceptible to participant deception, bias and misinterpretation. Epworth scores should not be viewed in isolation. Motivation to deceive such a subjective device may be high, especially where operational readiness could conceivably be measured by the ESS, such as in train engineers or truck drivers when stopped for a mid shift fatigue check. The ESS also measures self-reported behaviour in given situations which renders the device susceptible to recall bias. A further weakness in the situational paradigm is the availability of situational examples. One cannot fall asleep in a vehicle if one is never in one. Neither is it possible for most people to fall asleep after lunch while at work, whereas a retired person (no more or less objectively sleepy) may be able to do so given they have the opportunity.

In terms of concurrent validity with other measurements the ESS has a significant but relatively weak relationship the Multiple Sleep Latency Test (MSLT); the gold standard for assessment of daytime sleepiness (see also discussion on the MWT). Chervin and colleagues[232] find that the ESS-MSLT relationship is significant but weak (rho=-0.37, $p<0.005$). Kingshott and colleagues[90] report a similar relationship: ESS correlating significantly but not strongly with the MSLT (r=-0.23) and the MWT (r=-0.48). Benbadis and colleagues[259], by comparison, report no significant relationship.

Barnes and colleagues[6] report a surprisingly strong relationship between MSLT and ESS of R=0.50 in patients with mild OSAS. John's argues that the ESS measures a different facet of sleepiness than either the MSLT or the MWT, one more likely to reflect situational sleepiness in real life rather than a clinical or laboratory based setting[7, 8]. The reported relationships between the ESS and other sleepiness variables are not consistent. However the balance of evidence at this date point to a likely relationship between the gold standard MSLT and the ESS of around r=-0.20, -0.30 based on the findings of Kingshott et al. and Chervin et al.[3, 4]. The studies reported in chapters 2 3 and 4 have used both the Epworth Scale and the MWT because of the likely multifactorial nature of sleepiness.

The ESS was originally developed in English but has successfully been translated into German[9], Brazilian Portuguese[10], Polish[11], Chinese[12], and Spanish[13].

The ESS is often used as a clinical device to aid in the diagnosis of Excessive Daytime Sleepiness. However the diagnosis of a pathological condition requires the definition of a condition of normality as a basis. Johns'[1] first paper on this topic attempts to define normality with a sample of 30 hospital day-workers who did not report sleep disturbances (including snoring). In this sample an average of 5.9 with an SD of 2.2 was found within a range of 2-10. This is the first mention of the score of 10 as the upper reach of normal. Idiopathic Hypersomniacs and Narcoleptics, by comparison, scored between 12 and 24 with an average of about 17.7 and a SD of around 3.3. More equivocal were the findings in the groups diagnosed with other sleep disorders; OSAS (mean±SD; 11.7±4.6), Periodic Leg Movement Disorder (PLMD) (9.2±4.0), Primary Snoring (6.5±3.0), and Insomnia (2.2±2.0). Snorers, Insomniacs, and PLMD did not differ significantly from controls.

More importantly New Zealand population norms are now available for better comparison with clinical samples. Gander et al.[14] have found that about 15% of the population aged 30-60 had Epworth scores in excess of 10, but in Māori the figure was 23.7%. The average Epworth score in the non-Māori population is around 6.0 but in the indigenous population it is around 7.5. 5441 participants provided complete information for logistic modelling of the risk factors for excessive daytime sleepiness. Independent predictors of excessive sleepiness were being Māori, male, older, habitually sleeping

outside 7.5-8.0 hrs per night, never or rarely getting enough sleep, never or rarely waking refreshed, larger neck sizes (obesity marker), having observed breathing pauses during sleep (apnoea), and not drinking alcohol.

Clinical sample averages reported here in chapter 2 and 4 are clearly higher than average indicating substantial sleepiness. The studies reviewed in chapter 3 are likewise substantially higher than expected by comparison with the NZ population norm. These samples represent significantly sleepy people.

Internal consistency for the ESS is reported via Cronbach's alpha as being 0.88 in people with sleep disorders but borderline internal consistency of 0.73 in medical students[15]. Test retest reliability after 5 months within a group of Australian medical students was r=.82. Medical students in this study scored an average of 7.6±3.9; which was significantly lower in all eight questions than the 40 patients with OSAS, who scored on average 17.2±3.7. Johns' also compares self-rated ESS scores with ESS scores rated by the partners of patients. Spouses tended to rate their partners higher than those partners rated themselves. Overall this was not a significant overestimate but in 2 items spouses scored their partners significantly higher than self-ratings[7].

Johns'[16] also reports that in people with OSAS, as would be expected, as the Respiratory Disturbance Index (RDI) increases so to does the ESS. Together with minimum levels of overnight SaO2, RDI was used to predict ESS with R=0.460. Levels of RDI above 5 are generally regarded as evidence of sleep disordered breathing. As such those in a group with RDI<5 are usually regarded as a normative sample. Johns reports a group in this study with an RDI<5 but which snores (n=108) as having a mean ESS of 8.0(±3.5) ranging between 0-15. In contrast others have found no relationship between ESS and AHI. Kingshott and colleagues[4] report a non-significant negative correlation between AHI and ESS. This finding corroborated this groups finding that neither patient's nor their partners ratings of ESS in the patients correlated with AHI[17].The sleep heart health study[18] probably offers the best evidence of a relationship between the ESS and measures of nocturnal severity, in this case the RDI. Gottlieb and colleagues[18] report that in a middle aged population the proportion of people with ESS>10 increases from 21% in people with a non-pathological of RDI<5 to 35% in those with RDI>30 with a clear linear trend. Whilst this is a significant

relationship, it is also a weak one[18, 19]. The reason that RDI predicts so little of the variation in ESS are outlined by Gottlieb and colleagues[18], neither RDI nor ESS are perfect measures and RDI is not the only factor that might contribute to daytime sleepiness. This point is also discussed in chapter 3 as a possible contributor to the lack of effect of CPAP in alleviating sleepiness in patients with mild OSAS.

Jenkinson and colleagues[20] report that the ESS has only a modest relationship with the SF-36 summary scales in a clinical sample of people with OSAS (Correlation Coefficients; Physical –0.36 and Mental –0.28). Which is not dissimilar to the relationship reported between the ESS and other quality of life measures and between quality of life measures[20].

Briones et al.(Briones 1996 sleep) report that the ESS has a relationship to some of the subscales in the SF-36, specifically Vitality (r=-0.41), General health (r=-0.30), and role emotional (r=-0.30). In this same study the authors report a relationship of r=–0.27 between the ESS and the MSLT.

The ESS is also the most widely used measure of daytime improvement in RCTs of CPAP[6, 21-35]. In RCTs of CPAP in the treatment of OSAS the ESS has been demonstrated to sometimes be subject to a significant placebo effect[6, 26, 28, 29, 31, 32]. These studies show that ESS can improve even when inefficacious treatment is given and confirm the need to undertake controlled trials when investigating the effects of CPAP in OSAS. The size of the placebo effects on Epworth during CPAP trials warrants further investigation and should be quantified through a systematic search and meta-analysis. Those studies that investigated specifically mild OSAS (AHI 5-30) are shown in bold in the table and reviewed in more detail in Chapters 2 and 3.

Figure 1: Sensitivity of Epworth Scale to Treatment with CPAP in patients with OSAS

First Author and year	*Placebo condition improvement**	*Crossed over group improvement †*	*Parallel group Improvement‡*
Engleman 1997	**NS**	**NS**	
Engleman 1998	NS	p=0.001	
Engleman 1999	**NS**	**p=0.008**	
Fergusson 1997	§	NS	
Redline 1998	**§**		**NS**
Henke 2001	Sig.	NR	NS
Montaserio 2001	**§**		**NS**
Hack 2000	Sig.		p<0.001
Montserrat 2001	Sig.	p<0.001	p<0.001
Ballester 1999	§		p<0.001
Faccenda 2001	Sig.	p=0.001	
Jenkinson 1999	Sig.		p<0.001
Jokic 1999	§	NS	
Barnes, 2002	**Sig.**	**NS**	
Barnes 2004	**NS**	**p<0.001**	
Barbe, 2001	NS		NS

*The placebo condition exhibits significant improvement over time. † A crossover trial where CPAP arm improves significantly compared to the comparison treatment/placebo arm. ‡ Parallel trial where group receiving CPAP therapy shows significantly more improvement than the comparison treatment/placebo group. § No placebo condition as the alternative treatment is believed to be an active treatment - examples include conservative treatment, oral devices and positional therapy.
NR=Not Reported; NS=Not Significant (i.e. p>0.05). Studies shown in bold are specifically investigating mild OSAS with AHI<30.

1.2 MAINTENANCE OF WAKEFULNESS TEST

The Maintenance of Wakefulness Test (MWT) is related to it's predecessor the Multiple Sleep Latency Test (MSLT); both use polysomnographic measures- EEG, EOG, and EMG to measure the length of time it takes a person to go to sleep when left in a situation optimum for achieving a sleep state. The rationale for the MSLT in sleep medicine is that the drive to sleep will be highest in those with insufficient or fragmented sleep, all other things being equal. Commonly the MSLT gives 4-6 nap chances across the day beginning at 10am and spaced 2 hours apart[36, 37]. Instructions

to people doing the test are to "allow themselves to fall asleep" or something similar. Standardisation of the test is vital and guidelines have been published[37]. The MSLT is sensitive to experimental sleep restriction and to the effects of various sedating or alerting medications[38, 39] but does not seem to correlate well with the number of respiratory events per hour in people with sleep disordered breathing[40].

The MSLT is the gold standard for the objective measurement of daytime sleepiness[2]. However, is a test of daytime sleep propensity really relevant when the true measure of the success of a treatment for disorders of excessive sleepiness is arguably increased ability to stay awake? Staying awake is arguably more relevant an operational question.

The MWT is similar to the MSLT in that it objectively measures time to physiological sleep onset at multiple set times over the day. Two major differences exist; during the MWT rather than being reclined in a bed, as in the MSLT, patients are sitting up on the bed. The other major difference is that the instructions to the patients are to "remain awake" rather than "go to sleep" as in the MSLT[2]. Some researchers are of the opinion that these two tests, whilst methodologically similar, are most likely measuring separate facets of sleepiness[2, 41-43]. In the MWT it is failure to keep an activated awake system operating that unmasks sleep. The MSLT by comparison should immediately bypasses this wakefulness system to immediately test sleep propensity[2]. The key differences in the posture and motivation mean that the MWT measures alertness as well as the sleepiness that the MSLT measures[43]. It has been proposed that the MWT is the most relevant test of the two as a measure of operational fitness[41, 42].

Limited normative data for the MWT exist to delineate a gradient between abnormal and normal levels of sleepiness in the community. Doghramji and colleagues[44] have measured the MWT latencies for 64 healthy subjects (27 males 37 females) in the age ranges 30-69. Tests were carried out at two-hour intervals across the day, beginning at 10am, and a final test at 4pm. In this study four variations of the MWT were scored. The first two scoring variations were at a termination point of 20 minutes. Sleep onset was scored as either the first epoch of any sleep stage (MWT20) or at least three consecutive epochs of stage1 or any epoch of stage2 or deeper (SUSMWT20). Two forty minute length tests were also investigated using the same criteria (MWT40 and

SUSMWT40). Mean sleep latency for healthy males and females in the SUSMWT40 protocol were 36.2 (6.0) and 34.6 (9.0) respectively. These are currently the best available indicators of normality in the MWT.

Sangal and colleagues[41, 42] highlight an interesting phenomenon in sleep research of this type. After treatment with a variety of methods 26 patients with OSAS showed improvements in scores in the MWT but not in the MSLT, as part of a larger study. These authors are of the opinion that the MSLT and the MWT may measure different things. Treatment in this case affects the ability to stay awake (MWT) more than the ability to go to sleep quickly (MSLT), which remains intact. This finding has had some support from RCTs of CPAP in this area, including our own findings from chapter 3. The MWT and MSLT are differentially sensitive to treatment, at least with CPAP.

In another investigation by the same authors[41] the relationship between the MWT and MSLT was investigated in 258 consecutive patients referred for excessive sleepiness. Correlation between the tests was a significant r=0.41. Test retest reliability of the MWT over the day was r=.609. This figure is reasonably low in terms of such a measure but it must be remembered that the MWT is, in part, circadian dependent, so some variation is to be expected. The authors believe that two factors they name "Alertness" and "Sleepiness" explain most of the variance between the MSLT and MWT observed in patients of this type.

Mitler and colleagues[45] have further analysed the MWT data collected in the study by Doghramji and colleagues[44]. MWT latencies do not significantly correlate with gender, nocturnal sleep duration, or age. However, people in the age group 30-39 (30.9 mins ±8.9) were found to have significantly lower latencies than people in the sixties (38.0 ±3.7). There was however no linear finding of changing latencies across age groups, the observed difference in the 4th decade of life may be a chance finding[45].

In a small study of professional bus drivers, Häkkänen and colleagues[46] have shown an ability to discriminate between a group with mild OSAS (n=10) and one without SDB (n=10). Before CPAP treatment the mild OSAS group had an average MWT latency of 23.2 (10.2) compared with the non SDB group with a mean latency of 35.4 (6.8). This difference was significant (Wilcoxon 2 sample matched pairs, p=0.023).

After treatment with CPAP the mild OSAS group had significantly longer average latencies at 31.8 (6.03), a significant improvement from baseline (p=0.046). In terms of the present study the important facets of Häkkänen and colleagues[46] investigation are the ability of the MWT to discriminate between a mild sleep disordered group and another group without SDB. It is also demonstrated that the MWT is sensitive to treatment of mild OSAS with CPAP. It is not possible to tell whether some of the observed improvement is perhaps due to expectation of treatment and/or placebo effect as might be observed in a randomised controlled trial.

Poceta and colleagues[47] have presented the data from 322 patients with OSAS of various severities. Latencies on the MWT correlate significantly in the expected direction with some major indices of disease severity, respiratory arousal index (r=-.35), weight/height ratio (r=-.25), and mean oxygen saturation (r=.30). The latencies on the MWT do, with a large amount of variance, decrease as the severity level of OSAS increases. When the patients in this study are grouped into clinical disease severity steps according to respiratory arousal index a clear linear decrease in MWT latencies can be seen as the arousal index increases[47]. Poceta and colleagues present further data from 24 patients treated with CPAP which shows an improvement in MWT latencies from 18.0 (12.3) pre-treatment to 31.9 (10.4) post-treatment.

The MSLT[6, 21, 22, 25, 27, 48] and the MWT[23, 33], as well as a variation of the MWT called the OSLER[28, 32] have been used in randomised controlled trials of CPAP in the treatment of OSAS. The MSLT has however only been found to improve after CPAP in one trial[22]; a trial where the patients did not have mild OSAS (AHI>15, mean=43(37), n=23). The effects of CPAP on both measures in people with mild-moderate OSAS are meta-analysed in chapter 3.

Jokic and colleagues[33], in a crossover design, compared the effects of CPAP to positional therapy in a small group of patients (n=13) with positional OSAS. They found no benefits in terms of MWT due to CPAP compared to positional therapy designed to avoid sleeping supine. While the patients in this study had relatively mild OSAS (AHI 17.9 SEM 4.7) they were not recruited to specifically study the effects of CPAP on mild OSAS. Positional therapy can be effective in some cases, a lack of a positive finding in this study does not mean that CPAP is not better than placebo. It

does mean that in this subset of patients that a behavioural device should also be viewed as a treatment option, along with CPAP.

Engleman and colleagues[23] have presented a crossover study specifically dealing with people with mild OSAS (AHI5-15). They found that compared to an orally taken pill placebo CPAP offered no net benefit in terms of increasing MWT times (ES=0.19, $p<0.2$). However, patients in this trial were still regarded as pathologically sleepy by the authors in light of their MWT scores, even after CPAP treatment[23].

The next two studies presented were both based in Oxford, UK. This centre does not use the gold standard polysomnographic techniques, instead they rely upon oximetry to diagnose OSAS. The groups they have investigated tend to have what appears to be fairly severe OSAS, a situation which may explain their findings of improved latencies on the OSLER MWT. The relative severity of the OSAS may have left greater room between 'normal' and 'abnormal' scores: a gap large enough to be statistically significant after successful CPAP treatment. In the two studies above this may not have been possible due to the relatively mild OSAS treated, which left little room between 'normal' and 'mildly affected'. Alternatively the OSLER MWT itself may be more sensitive than the standardised MWT in detecting treatment response in people with OSAS. The OSLER also reflects behavioural performance rather than EEG criteria. It may be of more relevance to an operational setting where micro-sleeps may be just as important as standard MWT scored sleep for the arbitrary 30 or 90 second period. It is not possible at this stage to tell whether the observed differences were due to the severity of the OSAS or the device used or some combination of the two are responsible for the difference in findings shown.

The OSLER (Oxford SLEep Resistance test) is cheaper to administer and easier to score than the standard MWT, while still being reasonably accurate[49, 50]. Instead of using polysomnographic criteria for the establishment of sleep onset, this test uses behavioural criteria. Once every three seconds a light is activated, which must be responded to via a movement sensitive finger pad. If seven consecutive signals have been missed (i.e. 21 seconds) then the patient is regarded as being behaviourally asleep[49].

Jenkinson et al.[32] have published a parallel trial comparing CPAP (n=53) to subtherapeutic pressure CPAP (n=54). Having employed the 40 minute OSLER version of the MWT, they found significant a significant improvement on CPAP after comparison to placebo improvement.

Hack and colleagues[28] have also found a significant improvement in OSLER scores after treatment when CPAP group improvement (n=26) was compared with placebo sub-therapeutic pressure CPAP group improvement (n=33).

The MWT offers a valid objective measurement of sleepiness. It has been shown to reflect disease severity within a clinical group of people with OSAS[47]. It has also been shown to discriminate between a group of mild apnoeics and normal controls[46]. The MWT has been shown to be sensitive to change when people with OSAS are treated with CPAP[46, 47]. This last point is most contentious, RCTs of CPAP have not shown benefit in terms of MWT scores after treatment with CPAP compared to positional therapy[33] and tablet placebo. However benefit has been shown with a modified MWT, the OSLER[28, 32], which is not being used in this study. There is also a potentially worrying trend toward worsening in MWT latencies during CPAP RCTs that has been found in two recent RCTs by Barnes et al.[35] and that reported in chapter two. The measurement may not be stable over 4 iterations that are common in clinical crossover trials.

This study used a limited MWT polysomnographic montage with two EEG channels (C4-A1 and C3-A2), left and right EOG, and submentalis (chin) EMG as recommended by the newly updated guidelines but did not use single channel Electrocardiogram[51]. ECG is of uncertain utility in pinpointing sleep onset in an MWT. Patients were left alone in a darkened room with a stable temperature (20-24°C). They were deprived of time keeping devices, including cell-phones, and were instructed not to use motor behaviours, such as waving hands singing or slapping themselves, as methods of keeping awake. Participants were instructed to "Sit quietly and try to remain awake for as long as possible, please". This instruction was given to the participants, via intercom, immediately before the lights were turned out from the control room.

The version of the MWT employed in these two studies (chapters 2 and 4) used two tests per day rather than the more often used four per day. The two tests were carried out at 1:30pm and 3:30pm on each of the four days of patient visits in the mild OSAS study and 2 and 4pm in the C-flex study. Following the SUSMWT40 protocol of Doghramji and colleagues[44] sleep onset was established after 3 continuous epochs of stage 1 or any epoch of another sleep stage including REM. Time to sleep onset was then deemed to have begun from the first of these epochs of stage 1 or deeper. With only one exception the only other stage of sleep seen was stage 2.

The American Academy of Sleep Medicine has recently recommended that the MWT be employed with 4X40 minute naps throughout the day with the first trial beginning at about 9 or 10am[38, 51]. These recommendations were published after the trial reported here had begun and are thus not followed exactly. We have used the protocol described by Doghramji et al. as the best available at the time but with the possibly important modification of the number of trials used. Despite the breadth of evidence reviewed by the AASM task force there does not appear to be any justification for the use of as many sleep opportunities as recommended. The marginal benefit of 2 or 3 additional tests (compared to our two nap opportunity) does not appear to be discussed. As the test is so time consuming and onerous on patients I feel this is an important omission. What additional ability to discern abnormal daytime sleepiness do these additional tests outside the postprandial rise in sleep propensity give? This is unclear. Our position of using two tests is further strengthened because the timing of our naps is designed to be in the postprandial rise and is further defensible because it is a consistent measure within the trial: our results are internally consistent. As such it might not be appropriate in a clinical setting where one day measurements would need to be compared to the normative data available rather than to a patient's previous mean latency. The normative data available from Doghramji et al. and Banks et al. indicates no major differences between the mean latencies of any of the 4 naps[44, 52]. Given that a great deal of the subjects in these studies exhibited ceiling effects no additional benefit was gleaned from 2 additional tests where the subject did not fall asleep before the postprandial dip. Conversely we have employed two naps without much prior justification except an unpublished abstract that has not since been published in its full form[53]. We can justify this choice now by invoking patient comfort, internal consistency and the lack of

proof that four naps offers marginal benefit beyond two naps placed in the postprandial circadian increase in sleep propensity.

1.3 PSYCHOMOTOR VIGILANCE TASK

Simple reaction time can be measured via the Psychomotor Vigilance Task (PVT)[54]. The device is set to measure simple reaction time over ten minute samples. The physical device is a portable handheld box (20cm*11cm*5cm, weight 600g). The interface surface contains an LED above an LCD and screen and two buttons. The numeric display (LED) randomly indicates increasing reaction times (in milliseconds) and is terminated by, as fast as possible, pressing the left or right button (left for left handers, right button for right handers). Feedback on performance is given via the LED display, which pauses to display the reaction time for 1.5 seconds after each reaction. Each ten minute trial yielded approximately 80-100 RT values for analysis. The time between each reaction test varies randomly between 2 and 10 seconds. Failure to respond within 500 ms gives a recording of a 'Lapse'. Pressing the button before the LED screen begins counting up yields a 'False Start'. Respondents can also become worse on this test as the test goes on, this can be expressed as a slope decrement over time. Response times can also be characterised by the slowest and fastest 10% of scores and the median of the scores during a test. These studies have focussed on the treatment effects on mean, slowest 10% RTs and lapses.

Before and after each test the patient also responded to a PVT based question about their subjective sleepiness. A 10 point likert scale varied between 'No' and 'Yes' was given to answer the question posed, "Sleepy?". These metrics are very rarely reported however and were not analysed in these studies.

The environment during these tests was quiet with temperature (20-25°C) and ambient light steady. Patients were placed in a curtained but otherwise normally lit bedroom by themselves. The investigators were not present during the test. The patients were not interrupted during the test in order to avoid erroneous lapses that have been observed in operational settings[55] where the PVT has been employed.

Preceding each ten minute test was a 1 minute familiarisation or re-familiarisation test. This test was identical in all respects to the full ten minute test, save in terms of

duration. The object of this one minute test was to minimize any practice effect and to serve as a reorientation to the device for these somnolent patients. The data from these 1 minute tests is not kept or analysed.

The PVT has been reported to have the ability to discriminate between apnoea patients with RDI>10 and controls with RDI<10 who snore. Chugh and colleagues[56] presented findings from 22 untreated OSAS patients and 10 snoring controls. Controls had significantly less lapses, and faster scores on the 10% slowest reaction times than did OSAS patients. Controls also tended to have faster 10% fastest reaction times ($p=0.08$) and rate themselves as subjectively less sleepy after the test ($p=0.06$), although these did not reach significance[56].

Dinges and colleagues[57] have shown that PVT lapses increase in a dose response relationship across increasing clinical severity groups of RDI in commercial truck drivers. 198 truck drivers with suspected OSAS measured via a questionnaire based screening device were given PSG to determine RDI severity in groups from 0-5-15-30. Transformation of the number of lapses gave a significant increase across the severity groups ($F(3,193)$, $p=0.003$)[57].

The PVT has twice been used before in an RCT of CPAP in mild OSAS. Barnes and colleagues (2002) and Barnes and colleagues (2004) have both used the device. However in neither case was the reported mean reaction time improved due to CPAP therapy[6, 35].

1.4 HOSPITAL ANXIETY AND DEPRESSION SCALE

The HADS was developed by Zigmond and Snaith[58] and as it's name suggests the HADS screens for pathological anxiety and depression in clinical populations. A recent review of the HADS[59] concluded that the device has good internal consistency, concurrent validity, specificity and sensitivity. There is however an expected significant and strong correlation between anxiety and depression as measured by this device[59].

The HADS consists of 14 questions, 7 related to anxiety and 7 related to depression. Each of the 14 questions has 4 options that are scored between 0 and 3. Thus each of the scales 'Anxiety' and 'Depression' range between 0 and 21.

Mykletun and colleagues[60], based on a large community based sample (n=51,930) conclude that the scale is bi-dimensional (but intercorrelated) particularly in subsamples exhibiting symptoms of mental illness. Questions 6 and 7 may not be unique to Anxiety or Depression but seem to load on both. These authors also conclude that internal validity in the HADS, as measured by Cronbach's alpha, is satisfactory, being between 0.73 and 0.85 in all samples[60].

Developed to screen cheaply for mental disorders, is the HADS really sensitive enough to measure changes in mild OSAS patients? This group might be expected to be less likely to be clinically anxious or depressed than those whom the scale was originally validated? (any literature from mild OSAS overseas- descriptive).

The HADS had been used previously to Chapter 2 by only one research centre as an outcome measure in RCTs to measure the efficacy of CPAP in the treatment of mild[21, 23] and more severe[22, 48] OSAS. These methodologically similar studies have shown that the HADS depression scale is probably sensitive to changes bought about by 4 weeks of CPAP treatment in mild and more severe OSAS when compared to placebo[21, 23, 48]. HADS depression scores had been shown to significantly improve on arms involving active treatment in both trials of CPAP in mild OSAS[21, 23]. The HADS anxiety scale scores were better after the treatment arm than after the placebo arm in the group involved in the earliest study[48]. This finding has yet to be replicated and does not indicate improvement as no baseline measurement was taken. Interestingly in the later study[22] involving more severe OSAS, with AHI>15, CPAP therapy did not improve HADS scores when compared to placebo improvements.

Figure 2: Sensitivity of the Hospital Anxiety and Depression Scale to treatment with CPAP in patients with OSAS

HADS Subscale	*Anxiety*	*Depression*
Engleman et al. 1994	p=0.02	p=0.002
Engleman et al. 1997	**NS**	**p=0.03**
Engleman et al. 1998	NS	NS
Engleman et al. 1999	**NS**	**p=0.003**

Only p values less than 0.05 were considered significant. NS=Not Significant improvement when compared to placebo; Studies specifically investigating mild OSAS shown in bold.

1.5 SHORT FORM 36 QUESTION VERSION OF THE MEDICAL OUTCOMES SURVEY (SF-36)

The SF-36 is a generic health related quality of life measure. It is a widely validated 36 item abbreviated version (hence SF -short form) of the Medical outcomes survey (MOS) containing 8 independent sub scales[61-63]. As a general measure of subjective health the SF-36 aims to measure multiple dimensions of health that are generally applicable across disease groups. But this approach requires includes a trade-off between breadth and depth of enquiry and the expected time of completion of 5-10 minutes. General applicability saves some duplication of effort in the construction of disease specific measures of quality of life and more importantly allows rough comparisons across different disease groups of disease outcomes and effectiveness in terms of outcomes of any treatments within and across disease groups. Thus the changes are comparable with not only other OSAS afflicted people treated with different means but even with people being treated for altogether different diseases[61-63]. However, the Sf-36 does generally seem to be paired with a disease specific HQoL measure. We have also followed this pattern by combining the Functional Outcomes of Sleepiness Questionnaire with the Sf-36.

Each of the SF-36's subscales are scored between 0 and 100. These independent subscales are Physical functioning (PF-10 items), Role Physical (RP-4 items), Bodily pain (BP-2 items), General health perceptions (GH- 5 items), Vitality (VT-4 items), Social functioning (SF- 2 items), Role emotional (RE- 3 items), Mental health (MH- 5 items) and reported health change over a 1 year period (1 item)[62-65]. The frame of reference in the mild OSAS study was not 1 year, but two weeks- the length of the

washout period. These subscales primarily load onto and can be combined into Physical and Mental component summary scales[64]. However we did not do this.

The SF-36 appears to differentiate between the health status of people with or without sleep disordered breathing. Moore and colleagues[66] report that in a sample of 39 patients with moderate-severe OSAS that RDI, after controlling for gender and age, was significantly negatively correlated with mobility, social functioning, energy and fatigue and health distress. These subscales are part of the longer Medical Outcomes Survey and appear in abbreviated form in the SF-36. In the much larger Sleep Heart Health Study (n=5816), Baldwin and colleagues[67] used multiple logistic regression techniques to show that only the Vitality subscale of the SF36 has a linear, although weak, relationship over the 4 increasing categories of RDI severity (0-5-15-30-30+). That same study also showed that those with the most severe SDB (RDI 4% 30+) were more likely to report poor quality of life, after adjustment for age, gender, ethnicity, marital status, BMI, smoking, education, cardiovascular/respiratory conditions and use of sleeping pills, in terms of physical functioning, general health, vitality and social functioning than those with the least severe SDB (RDI<5)[67].

In the previous smaller community based Wisconsin Sleep Cohort study (n=737) Finn and colleagues[68] reported a linear relationship across all SF-36 subscales as AHI increased over a clinically relevant severity scale also used later[67]. Baldwin and colleagues[67] suggest this may be a statistical problem due to the non-continuous nature of some of the SF-36 subscales used in the linear regression employed by the Wisconsin group. Agreement was reached between the groups in terms of the reducing quality of life measured by the Vitality sub-scale as the severity of SDB increases[67, 68]. Thus even mild a group with SDB should demonstrate a difference from normal population averages given sufficient sample size. The sample reported in Chapter two is compared to population norms for New Zealand[69, 70] in that chapter and does show significant impairment in some of the subscales.

In a small investigation of people with mild OSAS (n=20, AHI 5-20) Gall and colleagues[71] report a impairment in quality of life when compared to people without SDB or daytime sleepiness (n=7) in some subscales of the SF-36 (Role Physical, Role Emotional, Social Functioning, Mental health, and Energy).

Yang and colleagues[72] have reported on the quality of life (SF-36) in a clinical sample of people with and without OSAS. In this sample, both the mild (AHI5-15, n=16) and the moderate-severe apnoeics (AHI>15, n=21), had impaired physical functioning and role-physical scores when compared to the group without SDB (AHI<5, n=46). The moderate-severe group had impaired Vitality when compared to the control group ($p<0.05$), but the mild group did not. The effects of OSAS may have been masked in this study by strong ceiling effects[72].

D'Ambrosio and colleagues[73], in a non-controlled trial of CPAP, showed that people with OSAS had significantly worse quality of life ($p<0.001$ in all 8 subscales except Bodily Pain) than normal people. After 8 weeks of CPAP treatment these researchers found the OSAS group had significantly worse HRQoL than normal people on only one subscale, Vitality[73]. Jenkinson, Stradling and Petersen[20] in a non-controlled trial found that the two SF-36 summary scores for Physical and Mental health were significantly worse than the general population at baseline but these improved significantly and were comparable to the general population after CPAP treatment. The authors concluded that the SF-36 is meaningful in an OSAS context and was sensitive to CPAP treatment in this group[20]. The severity of the OSAS in this group of patients was not reported.

Normative data for the New Zealand population are also available from the National Health Survey in 1996-7[69, 70] for comparison to our sample of patients with mild OSAS.

Problems may arise in the use of the SF-36 in that standard New Zealand scores are quite high compared to other countries and can approach the ceiling of 100 in some of the sub scales[69]. A baseline ceiling effect in the context of a treatment group in a randomised control trial this would result in the SF-36 having no sensitivity to improvements in quality of life resulting from CPAP treatment.

New Zealand normative values for the SF-36 suggest that gender, age, education, socio-economic status, and ethnicity differences exist[69, 70, 74]. In this crossover trial participants act as their own controls thus these slight demographic differences should

play little role unless they lead to ceiling effects. However as Scott and colleagues point out in older Maori and Pacific Islanders the clear two factor structure of the SF-36 becomes somewhat blurred[74]. Older Maori, in particular, exhibit a more unitary construct of health with mental and physical wellbeing not necessarily operating as separate constructs[74]. In the New Zealand context the scale has been shown to be sensitive to objective measures of health such as smoking status and hypertension, as well as the pseudo-objective indicator of GP visit rates[69]. The SF-36 is a reliable and valid measure of HRQoL in the New Zealand context[70].

The SF-36 has been employed in previous randomised trials of CPAP in the treatment of mild[23, 25] and more severe OSAS[29, 32] to assess improvements in HRQoL.

Redline and colleagues[25] found that the vitality subscale of the SF-36 was the only measure employed which had significant improvement when CPAP treatment was compared to Conservative medical treatment in their 8 week parallel mild OSAS groups (RDI 5-30).

Engleman and colleagues[23] investigated a mild sample of OSAS patients (AHI 5-15). The patients were randomised to begin either effective CPAP treatment or an oral placebo, after 4 weeks these conditions were crossed over. These investigators used the SF-36 as one of a number of measures of wellbeing and health status. Five of the nine SF-36 sub-scales showed improvements compared to placebo, they were; 'Health transition', 'Role-Physical', 'Bodily Pain', 'Social function', and 'Vitality'.

Jenkinson and colleagues[32] reported a 4 week parallel trial of CPAP vs. sub-therapeutic CPAP in a group of OSAS patients diagnosed with greater than 10 per hour blood oxygen dips (4% or greater) and scores greater than 10 on the Epworth sleepiness scale. SF-36 scores in Vitality, General Health Perception, and both the Mental and Physical component summaries were significantly improved when effective CPAP was compared to CPAP at sub-therapeutic pressure.

Montserrat and colleagues[29] employed a partial crossover design where one group received sub therapeutic pressure CPAP for half the trial followed by effective CPAP, a

crossover. The second group received CPAP throughout the trial. The proper comparison should thus be parallel and made between the groups, as the crossover group completes the sub-therapeutic arm group of the trial, and within the crossed over group after the CPAP arm of the trial. The patients in this trial had severe OSAwith associated daytime sleepiness, mean (standard deviation); AHI 54 (19) and ESS 16 (5). In the between groups comparison no measure in the SF-36 showed improvement over the six week period. When the crossover group was compared to itself effective CPAP treatment was significantly more effective in improving 3 of the 10 scales and sub scales; they were Bodily Pain, Social Functioning, and Role Emotional.

McFadyen and colleagues[34] present a non-randomised parallel trial with allocation based on placement on a waiting list. Patients who received fully titrated CPAP treatment were compared to a group receiving conservative treatments including weight loss, sleeping position, and alcohol avoidance. Both patient groups had non-different severity measures (CPAP group; AHI 46±40, 23%♀, n=44- Conservative group; AHI 39±16, 24%♀, n=25). After 3 months treatment the CPAP groups had significantly improved on all subscales of the SF-36 when compared to the conservative measures group. All of these improvements were also clinically significant with effect sizes being large (ES>0.80, Physical Functioning, Role Physical, Role Emotional, Mental health, Vitality) and moderate (0.4<ES<0.79, Social functioning, Bodily Pain, General Health). By far the largest improvement was observed in the Vitality subscale (ES=2.26). Both mental and physical summary scores in the CPAP group also showed large improvements when compared to the CT group[34]. These effect sizes are mitigated by the significant decrease in reported quality of life of the CT group. Thus the effect sizes for improvement within the CPAP group are actually lower than comparison of improvements between the groups. The seemingly non- randomised nature of this trial does however reduce it's strength in terms of other stronger trials undertaken.

Barnes and colleagues[6] have reported a randomised crossover trial using an oral placebo. The patients were 44 Australians in two centres with mild (AHI 5-30) obstructive sleep apnoea, however only 28 of these patients completed the protocol. In this study no significant effect of CPAP was reported after comparing the improvements on placebo and on active treatment. A placebo effect, where scores improved compared to baseline on placebo tablets, was observed in 4 of the 8 subscales. General quality of

life did improve over the trial but CPAP was not more effective than placebo in elevating HRQoL[6].

Figure 3: Sensitivity of the SF-36 to treatment with CPAP in patients with OSAS

SF-36 Scale	*PC*	*MC*	*PF*	*RP*	*BP*	*GH*	*VT*	*SF*	*RE*	*MH*
Redline, 1998	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**	**.05**	**NS**	**NS**	**NS**
Engleman, 1999			**NS**	**.03**	**.02**	**NS**	**.001**	**.01**	**NS**	**NS**
Jenkinson, 1999	NS	.002	NS	NS	NS	.002	.001	NS	NS	NS
Montserrat 2001 - Within group	NS	NS	NS	NS	.007	NS	NS	.024	.025	NS
- Between groups	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Barnes, 2002	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**
Barnes, 2004- 0.05										

NS=Not Significant; PC=Physical component summary, MC= Mental component summary, PF=Physical functioning, RP= Role Physical, BP= Bodily Pain, GH= General Health, VT=Vitality, SF= Social functioning, RE= Role emotional, MH= Mental health. Decimal numbers in cells represent p values less than the indicated figure -only p values below 0.05 were considered significant. Studies specifically investigating mild OSAS are shown in bold. Barnes et al. 2004 reported only the mean SF-36 score which was improved by 2.9 points after placebo adjustment

The use of the SF-36 is not without some real issues. Health status is measured via 8 scales that are assumed to be of equal importance to people. It may be the case that in people with OSAS, or New Zealanders in general, that only some factors measured by the SF-36 are valued as being important[69]. Thus improvement (or non-improvement) in one or few of the subscales of the SF-36 may be enough to warrant CPAP as an effective treatment. The reverse is also true: demonstrated improvements on measures deemed to be unimportant by patients do not warrant the general acceptance of a treatment. This measure also doesn't gauge sleep adequacy[63]. This factor was eliminated for brevity and due to it's relation with other subscales that are still found in the SF-36. One might, however, expect a measure of sleep adequacy, to be particularly sensitive in this study. In this trial such an allowance has been made with the inclusion of the Functional Outcomes of Sleep Questionnaire (FOSQ)[75].

1.6 FUNCTIONAL OUTCOMES OF SLEEPINESS QUESTIONNAIRE

The Functional Outcomes of Sleep Questionnaire (FOSQ) was designed to assess the impact on peoples' lives of disorders of excessive sleepiness[75]. It indicates a move away from using measures such as the Epworth Sleepiness Scale or the Stanford Sleepiness Scale to quantify improvement post treatment. The FOSQ attempts to measure the impact of disorders of excessive daytime sleepiness on day to day living. In this way the FOSQ may be more salient to Sleep Disorders Medicine than the SF-36 but does not allow the range of comparison across disease groups. The FOSQ consists of 5 sub scales ('General productivity'- 8 items, 'Social outcome'- 2 items, 'Activity Level'- 9 items, 'Vigilance'- 7 items, and 'Intimacy and Sexual Relationship'-4 items). Because these subscales contain different numbers of items the subscale score is derived by expressing each subscale total at a value between I and 4. In this trial the questions pertaining to intimacy and sexual relationships have been discarded as they require patient's to be in such a relationship. Retaining these questions may have thus required the exclusion from the trial of patients who make up a representative sample of those in the Wellington clinical population, an avoidable source of bias. To retain comparability of function with other studies the 4-16 points 4 subscale FOSQ scale can be adjusted to give a score out of 20.

Weaver and colleagues[75] have demonstrated that the FOSQ is both a reliable and valid measure of sleep related quality of life in three samples of clinical referrals who were then diagnosed as having OSAS or not having OSAS (n1=153, n2=24, and n3=51). Internal validity was demonstrated through inter item correlation (Cronbach's alpha) as being 0.95 for the full FOSQ. Reliability has been demonstrated through test-retest correlations of r=0.90 for the full FOSQ and between r=.81 and r=0.90 for the subscales.

Concurrent reliability has been established between the SF-36 and the Sickness Impact Profile (SIP). Global FOSQ scores are correlated with SF-36 Bodily pain and Role emotional. The FOSQ subscales are correlated with some of the SF-36 subscales except for Vigilance and Intimacy subscales that have not yet demonstrated a significant relationship[75]. The FOSQ is related to the SF-36 in some important areas but also seems to have added factors that may be more sensitive to changes in diseases of

excessive sleepiness. In this study the FOSQ was able to discriminate between people with OSAS (n=133) and those without (n=20) in a between groups analysis. The specificity and sensitivity of this measure in predicting individual disability is unclear. It is likely that further validation may have to include larger numbers of normal people without OSAS and who have not been clinically referred to a sleep disorders centre but have been PSG tested and do not have OSAS.

The FOSQ has been employed in two recent RCTs of CPAP in the treatment of severe[29] and mild[27] OSAS. Montserrat and colleagues[29] using a partial crossover design found no change in the FOSQ in the group receiving sub-therapeutic CPAP and then CPAP. These researchers did find a difference in the parallel part of that trial with General productivity and Vigilance both improving more in the CPAP group than the placebo group[29]. In three of the five subscales and in the total summary score the FOSQ was found to significantly improve during the sub-therapeutic CPAP phase of the trial (Social Outcome, Activity Level, Vigilance, and total FOSQ all $p<0.02$). This indicates that the FOSQ is sensitive to a placebo effect and should not be compared without a control.

Monasterio and colleagues[27] used a parallel trial over six months to show that CPAP+CT treatment of mild OSAS (AHI 10-30) was not better than conservative medical treatment alone (CT) in terms of bettering scores on the FOSQ. This study did indicate a trend toward CPAP+CT being more effective than CT as the difference in improvements from baseline between the groups after 6 months approached significance ($p=0.06$). None of the FOSQ subscales were found to significantly improve, although Vigilance also approached significance ($p=0.06$).

McFadyen and colleagues[34] using a non-randomised parallel trial of CT vs. CPAP in the treatment of a clinical population of people with OSAS have shown CPAP to be more effective than CT. Effect size analysis has shown in this study that CPAP has a large clinical effect ($ES<0.80$) in all four FOSQ subscales used (sexual functioning subscale was not used in this study) and the total FOSQ summary score when compared to CT. The reasons for these large effect sizes are briefly discussed in the SF-36 literature discussion, the same effects is also observed in the FOSQ subscales.

The study presented by Barnes and co-authors[6] shows significant improvements in the placebo arm of a crossover trial involving patients with mild OSAS. No significant net improvements were found when improvement on CPAP was compared to improvement on oral placebo, however vigilance did approach significance (p=0.06).

Figure 4: Sensitivity of the Functional Outcomes of Sleepiness Questionnaire to treatment with CPAP in patients with OSAS

FOSQ Subscales	*Gen.*	*Soc.*	*Act.*	*Vig.*	*ISR*	*FOSQ Total*
Montaserio, 2001	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**
Montserrat et al., 2001-						
Between groups	0.044	NS	NS	0.009	NS	NS
-Within crossover group	NS	NS	NS	NS	NS	NS
McFadyen, 2001	0.001	0.001	0.001	0.001	NR	0.001
Barnes, 2002	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**
Barnes, 2004						**NS**

Gen.- General Productivity; Soc.- Social Outcomes; Act.- Activity Level; Vig.- Vigilance; ISR- Intimacy and sexual relations. NS=Not significant. NR= Not reported. Decimals represent p values less than the indicted figure. Only p values less than 0.05 are deemed significant. McFadyen and colleagues[34] presented an Effect Size analysis on non-random parallel groups. Studies specifically investigating mild OSAS shown in bold.

REFERENCES: APPENDIX 1

1. Johns MW, A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*, 1991. **14**: 540-45.
2. Mitler MM, Carskadon MA, and Hirshkowitz M, *Evaluating sleepiness*, in *Principles and practice of sleep medicine (3rd ed.)*, MH Kryger, T Roth, and WC Dement, Editors. 2000, WB Saunders Co.: USA.
3. Chervin R, Aldrich M, Pickett R, and Guilleminault C, Comparison of the results of the epworth sleepiness scale and the multiple sleep latency test. *J Psychosom Res*, 1997. **42**(2): 145-55.
4. Kingshott RN, Engleman HM, Deary IJ, and Douglas NJ, Does arousal frequency predict daytime function? *Eur Respir J*, 1998. **12**: 1264-70.
5. Benbadis S, Mascha E, Perry M, Wolgamuth B, and Smolley L, Association between the Epworth Sleepiness Scale and the Multiple Sleep Latency Test in a clinical population. *Arch Int Med*, 1999. **130**(4): 289-92.
6. Barnes M, Houston D, Worsnop C, Neill A, Mykytyn I, Kay A, Trindler J, Saunders N, McEvoy R, and Pierce R, A randomized controlled trial continuous positive airway pressure in mild obstructive apnea. *Am J Respir Crit Care Med*, 2002. **165**: 773-80.
7. Johns MW, Sleepiness in different situations measured by the Epworth sleepiness scale. *Sleep*, 1994. **17**(8): 703-10.
8. Johns M, Reply to RD Chervin- The multiple sleep latency test and Epworth sleepiness scale in the assessment of daytime sleepiness (letter). *J Sleep Res*, 2000. **9**(4): 400-01.
9. Bloch K, Schoch O, Zhang J, and Russi E, German version of the Epworth sleepiness scale. *Respiration*, 1999. **66**: 440-47.
10. Alameddine M, RC A, Pedroso A, Aloe F, Tavares S, and AB S, Epworth sleepiness scale outcome in 181 Brazilian adults (abstract). *Sleep*, 1999. **22**(suppl. #1): S79.
11. Zielinski J, Zgierska A, Polakowska M, Finn L, Kurjata P, Kupsc W, and al. e, Snoring and excessive daytime sleepiness among Polish middle aged adults. *Eur Respir J*, 1999. **14**: 946-50.
12. Chung K, Use of the Epworth sleepiness scale in Chinese patients with obstructive sleep apnea and normal hospital employees. *J Psychsom Res*, 2000. **49**: 367-72.
13. Izquieredo-Vicario Y, Ramos-Platon M-J, Conesa-Peraleja D, and Lozano-Parra A, Epworth sleepiness scale in a sample of the Spanish population (letter). *Sleep*, 1997. **20**(8): 676-77.
14. Gander P, Marshall N, Harris R, and Reid P, The Epworth Sleepiness Scale: Influence of age, ethnicity, and socio-economic deprivation. *Sleep*, 2005. **28**(2): 249-53.
15. Johns MW, Reliability and factor analysis of the Epworth sleepiness scale. *Sleep*, 1992. **15**(4): 376-81.
16. Johns MW, Daytime sleepiness, snoring, and obstructive sleep apnea: The Epworth sleepiness scale. *Chest*, 1993. **103**: 30-36.
17. Kingshott R, Sime P, Engleman H, and Douglas N, Self assessment of daytime sleepiness: patient versus partner. *Thorax*, 1995. **50**: 994-95.
18. Gottlieb D, Whitney C, Bonekat W, Iber C, James G, Lebowitz M, and al. e, Relation of sleepiness to respiratory disturbance index: the sleep heart health study. *Am J Respir Crit Care Med*, 1999. **159**: 502-07.

19. Chervin R, The multiple sleep latency test and Epworth sleepiness scale in the assessment of daytime sleepiness (letter). *J Sleep Res*, 2000. **9**(4): 399-401.
20. Jenkinson C, Stradling J, and Petersen S, How should we evaluate health status? A comparison of three methods in patients presenting with obstructive sleep apnoea. *Qual Life Res*, 1998. **7**: 95-100.
21. Engleman H, Martin S, Deary I, and Douglas N, Effect of CPAP therapy on daytime function in patients with mild sleep apnoea/hypopnoea syndrome. *Thorax*, 1997. **52**: 114-19.
22. Engleman H, Martin S, Kingshott R, Mackay T, Deary I, and Douglas N, Randomised placebo controlled trial of daytime function after continuous positive airway pressure (CPAP) therapy for the sleep apnoea/hypopnoea syndrome. *Thorax*, 1998. **53**: 341-45.
23. Engleman H, Kingshott R, Wraith P, Mackay T, Deary I, and Douglas N, Randomised placebo controlled crossover trial of continuous positive airway pressure for mild sleep apnea/hypopnea syndrome. *Am J Respir Crit Care Med*, 1999. **159**: 461-67.
24. Fergusson K, Ono T, Lowe A, Al-Majed S, Love L, and Fleetham J, A short term controlled trial of an adjustable oral appliance for the treatment of mild obstructive sleep apnoea. *Thorax*, 1997. **52**: 362-68.
25. Redline S, Adams N, Strauss ME, Roebuck T, Winters M, and Rosenburg C, Improvement of mild sleep disordered breathing with CPAP compared with conservative therapy. *Am J Respir Crit Care Med*, 1998. **157**: 858-65.
26. Henke KG, Grady JJ, and Kuna ST, Effect of nasal continuous positive airway pressure on neuropsychological function in sleep apnea-hypopnea syndrome: a randomized placebo-controlled trial. *Am J Respir Crit Care Med*, 2001. **163**: 911-17.
27. Monasterio C, Vidal S, Duran J, Ferrer M, Carmona C, Barbe F, Mayos M, Gonzalez-Mangado N, Juncadella M, Navarro A, Barreira R, Capote F, Mayoralas L, Peces-Barba G, Alonso J, and Montserrat J, Effectiveness of continuous positive airway pressure in mild sleep apnea-hypopnea syndrome. *Am J Respir Crit Care Med*, 2001. **164**: 939-43.
28. Hack M, Davies RJO, Mullins R, Choi SJ, Ramdassingh-Dow S, Jenkinson C, and Stradling JR, Randomised prospective trial of therapeutic vs. sub-therapeutic nasal continuous airway pressure on simulated steering performance in patients with obstructive sleep apnoea. *Thorax*, 2000. **55**: 224-31.
29. Montserrat JM, Ferrer M, Hernandez L, Farre R, Vilagut G, Navajas D, Badia JR, Carrasco E, de Pablo J, and Ballester E, Effectiveness of CPAP treatment in daytime function in sleep apnea syndrome: a randomized controlled study with an optimized placebo. *Am J Respir Crit Care Med*, 2001. **164**: 608-13.
30. Ballester E, Badia J, Hernandez L, Carasco E, de Pablo J, Fornas C, Rodriquez-Roisin R, and Montserrat J, Evidence of the effectiveness of continuous positive airway pressure in the treatment of sleep apnea/hypopnea syndrome. *Am J Respir Crit Care Med*, 1999. **159**: 495-501.
31. Faccenda J, Mackay T, Boon N, and Douglas N, Randomized placebo-controlled trial of continuous positive airway pressure on blood pressure in the sleep apnea-hypopnea syndrome. *Am J Respir Crit Care Med*, 2001. **163**: 344-48.
32. Jenkinson C, Davies RJO, Mullins R, and Stradling JR, Comparison of therapeutic and subtherapeutic nasal continuous positive airway pressure for obstructive sleep apnoea: a randomised prospective parallel trial. *Lancet*, 1999. **353**: 2100-05.

33. Jokic R, Klimaszewski A, Crossley M, Sridhar G, and Fitzpatrick M, Positional treatment vs continuous positive airway pressure in patients with positional obstructive sleep apnea syndrome. *Chest*, 1999. **115**(3): 771-81.
34. McFadyen T, Espie C, McArdle N, Douglas N, and Engleman H, Controlled prospective trial of psychosocial function before and after continuous positive airway pressure therapy. *Eur Respir J*, 2001. **18**(6): 996-1002.
35. Barnes M, McEvoy RD, Banks S, Tarquinio N, Murray CG, Vowles N, and Pierce RJ, Efficacy of Positive Airway Pressure and Oral Appliance in Mild to Moderate Obstructive Sleep Apnea. *Am. J. Respir. Crit. Care Med.*, 2004. **170**(6): 656-64.
36. Carskadon M and Dement W, The multiple sleep latency test: what does it measure? *Sleep*, 1982. **5**: S67-S72.
37. Carskadon M, Dement W, Mitler M, Roth T, Westbrook P, and Keenan S, Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness. *Sleep*, 1986. **9**: 519-24.
38. Arand D, Bonnet M, Hurwitz T, Mittler M, Rosa R, and Sangal R, The clincal use of the MSLT and MWT. *Sleep*, 2004. **28**(1): 123-44.
39. Carskadon M and Dement W, Cumulative effects of sleep restriction on daytime sleepiness. *Psychophysiology*, 1981. **18**: 107-13.
40. Chervin R, Kraemer H, and Guilleminault C, Correlates of sleep latency on the multiple sleep latency test in a clinical population. *Electroenceph Clin Neurophys*, 1995. **95**: 147-53.
41. Sangal R, Thomas L, and Mitler MM, Maintenance of wakefulness test and multiple sleep latency test: measurement of different abilities in patients with sleep disorders. *Clin Neurophysiol*, 1992. **110**(1999): 2131-35.
42. Sangal R, Thomas L, and Mitler MM, Disorders of excessive sleepiness: treatment improves ability to stay awake but does not reduce sleepiness. *Chest*, 1992. **102**(3): 699-703.
43. Bonnet MH and Arand D, Arousal components which differentiate the MWT from the MSLT. *Sleep*, 2001. **24**(4): 441-47.
44. Doghramji K, Mitler M, Sangal R, Shapiro C, Taylor S, Walsleben J, Belisle C, Erman M, Hayduk R, Hosn R, O'Malley E, Sangal J, Schutte S, and Youakim J, A normative study of the maintenance of wakefulness test (MWT). *Electroenceph Clin Neurophys*, 1997. **103**: 554-62.
45. Mitler MM, Doghramji K, and Shapiro C, The Maintenance of Wakefulness test: Normative data by age. *J Psychsom Res*, 2000. **49**: 363-65.
46. Hakkanen H, Summala H, Partinen M, Tiihonen M, and Silvo J, Blink duration as an indicator of driver sleepiness in professional bus drivers. *Sleep*, 1999. **22**(6): 798-802.
47. Poceta JS, Timms RM, Jeong D-U, Ho S-L, Erman MK, and Mitler MM, Maintenance of wakefulness test in obstructive sleep apnea syndrome. *Chest*, 1992. **101**(4): 893-97.
48. Engleman H, Martin S, Deary I, and Douglas N, Effect of nasal continuous positive airway pressure treatment on daytime function in sleep apnoea/hypopnoea syndrome. *Lancet*, 1994. **343**: 572-75.
49. Bennet L, Stradling J, and Davies R, A behavioural test to assess daytime sleepiness in obstructive sleep apnoea. *Journal of Sleep Research*, 1997. **6**: 142-45.
50. Priest B, Brichard C, Aubert G, Liistro G, and Rodenstein DO, Microsleep during a simplified maintenance of wakefulness test: a validation of the OSLER test. *Am J Respir Crit Care Med*, 2001. **163**: 1619-25.

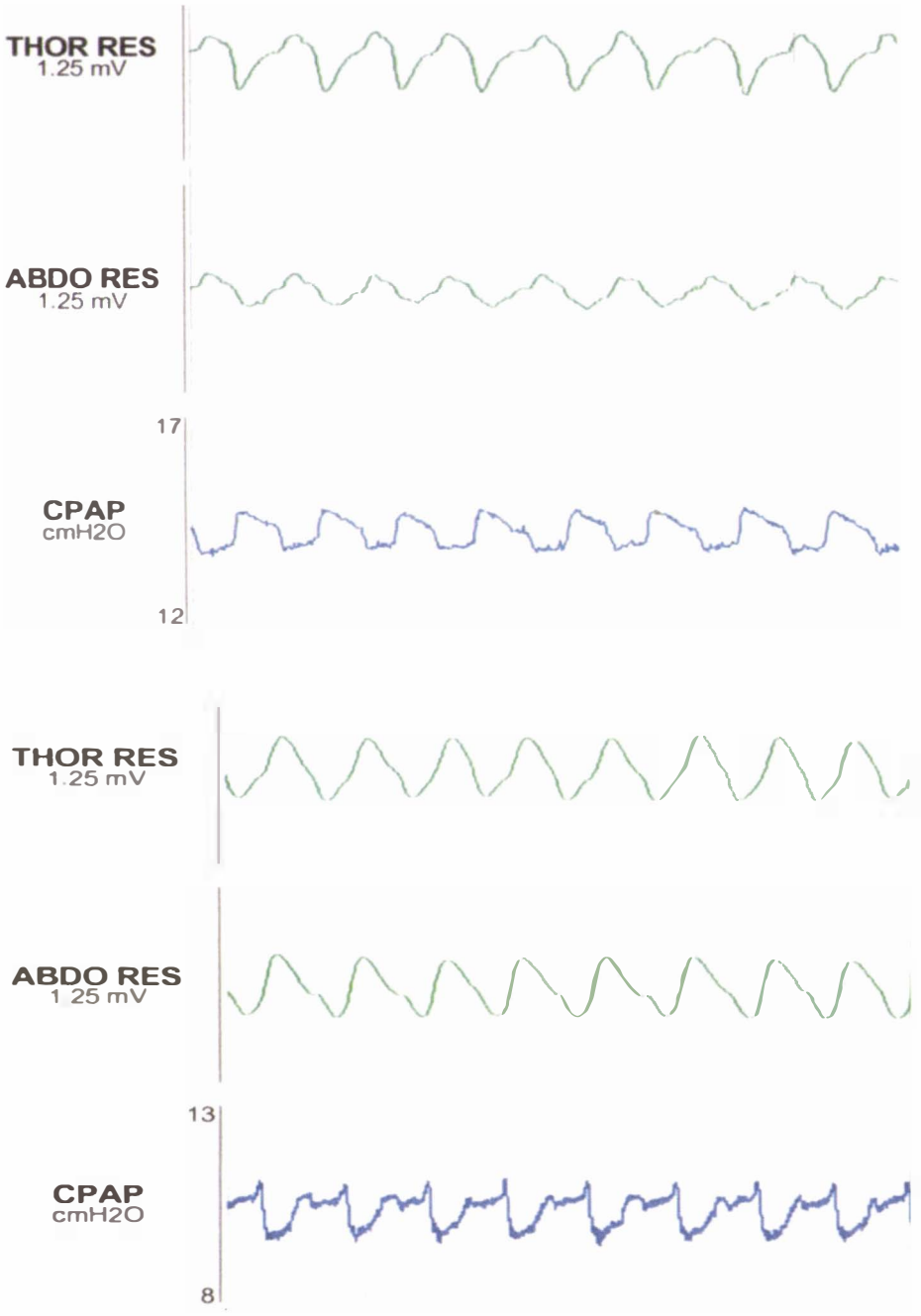
51. Littner M, Kushida C, Wise M, Davila D, Morgenthaler T, Lee-Chiong T, Hirshkowitz M, Loube D, Bailey D, Berry R, Kapen S, and Kramer M, Practice parameters forthe clinical use of the multiple sleep latency test and the maintenance of wakefulness test. *Sleep*, 2004. **28**(1): 113-21.
52. Banks S, Barnes M, Tarquinio N, Pierce R, Lack L, and McEvoy R, The maintenance of wakefulness test in normal healthy subjects. *Sleep*, 2004. **27**(4): 799-802.
53. Banks S, Kenner A, Barnes M, Pierce R, and McEvoy R. *Can the maintenance of wakefulness test be shortened? A preliminary investigation.* in *Annual Scientific Meeting of the Australasian Sleep Association.* 2000. Melbourne Australia.
54. Dinges DF and Powell NB, Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behav Res Methods Instrum Comput*, 1985. **17**(6): 652-55.
55. Miller M, *Hours of Work, Sleep Loss and Performance of Anaesthetists: Implications for Safety in Medical Practice*, in *Department of Public Health.* 2001, University of Otago: Dunedin.
56. Chugh DK, Weaver TE, and Dinges DF, Psychomotor vigilance performance in sleep apnea patients compared to patients presenting with snoring without apnea. *Sleep*, 1998. **21**(S1): 491.K1.
57. Dinges DF, Maislin G, Staley B, Pack F, Woodle C, and Pack AI, Sleepiness and neurobehavioral functioning in relation to apnea severity in a cohort of commercial truck drivers. *Sleep*, 1998. **21**(S1): 199.K1.
58. Zigmond A and Snaith R, The hospital anxiety and depression scale. *Acta Psychiatr Scan*, 1983. **67**: 367-70.
59. Bjelland I, Dahl A, Haug T, and Neckelman D, The validity of the Hospital Anxiety and Depression scale: an updated literature review. *J Psychosom Res*, 2002. **52**: 69-77.
60. Mykletun A, Stordal E, and Dahl A, Hospital anxiety and depression scale: factor structure, item analyses and internal consistency in a large population. *Brit J Psych*, 2001. **179**(6): 540-49.
61. Stewart A, Hays R, and Ware J, The MOS short-form general health survey: reliability and validity in a patient population. *Med Care*, 1988. **26**(7): 724-35.
62. Ware J and Sherbourne C, The MOS 36-item short-form health survey (SF-36): Conceptual framework and item selection. *Med Care*, 1992. **30**(6): 473-83.
63. Ware J and Gandek B, Overview of the SF-36 health survey and the international quality of life assessment (IQOLA) project. *J Clin Epidemiol*, 1998. **51**(11): 903-12.
64. Ware J, Kosinski M, Gandek B, Aaronson N, Apolone G, Bech P, Brazier J, Bullinger M, Kaasa S, Leplege A, Prieto L, and Sullivan M, The factor structure of the SF-36 health survey in 10 countries: results from the IQOLA project. *J Clin Epidemiol*, 1998. **51**(11): 1159-65.
65. McHorney CA, Ware J, Lu JF, and Sherbourne CD, The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care*, 1994. **32**(1): 40-66.
66. Moore P, Bardwell W, Ancoli-Israel S, and Dimsdale J, Association between polysomnographic sleep measures and health related quality of life in obstructive sleep apnea. *J Sleep Res*, 2001. **10**: 303-08.
67. Baldwin C, Griffith K, Nieto F, O'Connor G, Walsleben J, and Redline S, The association of sleep disordered breathing and sleep symptoms with quality of life in the sleep heart health study. *Sleep*, 2001. **24**(1): 96-105.

68. Finn L, Young T, Palta M, and Fryback D, Sleep disordered breathing and self-reported general health status in the Wisconsin sleep cohort study. *Sleep*, 1998. **21**(7): 701-06.
69. New Zealand Ministry of Health, *Our health, our future: the health of New Zealanders*. 1999, Ministry of Health: Wellington. 444.
70. Scott K, Tobias MI, Sarfati D, and SJ H, SF-36 health survey reliability, validity, and norms for New Zealand. *Aust NZ J Pub Heal*, 1999. **23**: 401-06.
71. Gall R, Isaac L, and Kryger M, Quality of life in mild obstructive sleep apnea. *Sleep*, 1993. **16**: S59-S61.
72. Yang E, Hla M, McHorney CA, Havighurst T, Badr S, and Weber S, Sleep apnea and quality of life. *Sleep*, 2000. **23**(4): 535-41.
73. D'Ambrosio C, Bowman T, and Mohsenin V, Quality of life in patients with obstructive sleep apnea: effect of nasal continuous positive airway pressure- a prospective study. *Chest*, 1999. **115**: 123-29.
74. Scott K, Sarfati D, Tobias M, and Haslett S, A challenge to the cross-cultural validity of the SF36 health survey: factor structure in Maori, Pacific, And New Zealand european ethnic groups. *Soc Sci Med*, 2000. **51**: 1655-64.
75. Weaver TE, Laizner AM, Evans LK, Maislin G, Chugh DK, Lyon K, Smith PL, Schwartz AR, Redline S, Pack AI, and Dinges DF, An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep*, 1997. **20**(10): 835-43.

APPENDIX 2
PUBLISHED PAPERS

APPENDIX 3
SLEEP QUESTIONNAIRE

APPENDIX 4
C-FLEX AND CPAP POLYSOMNOGRAPHY TRACES



C-Flex trace (immediately above) indicates pressure inside the mask that is different from CPAP (top). Typical curve at the top of exhalation is attenuated.

Erratum: Chapter 4

During the final collation of this thesis the citations in Chapter 4 failed to automatically convert into the Reference section. Unfortunately, this error was missed until after the thesis had already been submitted in its final form. As such the citations in Chapter 4 are still in the original Endnote citation code. Fortunately this can be read in an (Author, date) form as no two of the citations are from the same author. The reference section for Chapter 4 is supplied below.

Chapter 4 References

Aloia, MS, Stanchina, M, Arnedt, JT, Malhotra, A and Millman, RP (2005). Treatment adherence and outcomes in flexible vs standard continuous positive airway pressure therapy. Chest **127**(6): 2085-2093.

American Sleep Disorders Association (1992). EEG arousals: Scoring rules and examples. Sleep **15**: 173-184.

Ayas, N, Patel, S, Malhotra, A, Schulzer, M, Malhotra, M, Jung, D, Fleetham, JA and White, D (2004). Auto-titrating versus standard continuous positive airway pressure for the treatment of obstructive sleep apnea: Results of a meta-analysis. Sleep **27**(2): 249-253.

Banks, S, Barnes, M, Tarquinio, N, Pierce, R, Lack, L and McEvoy, R (2004). The Maintenance of Wakefulness Test in normal healthy subjects. Sleep **27**(4): 799-802.

Berry, R, Parish, J and Hartse, K (2002). The use of auto-titrating continuous positive airway pressure for treatment of adult obstructive sleep apnea. Sleep **25**(2): 148-159.

Dinges, DF and Powell, NB (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. Behav Res Methods Instrum Comput **17**(6): 652-655.

Doghramji, K, Mitler, M, Sangal, R, Shapiro, C, Taylor, S, Walsleben, J, Belisle, C, Erman, M, Hayduk, R, Hosn, R, O'Malley, E, Sangal, J, Schutte, S and Youakim, J (1997). A normative study of the Maintenance of Wakefulness Test (MWT). *Electroenceph Clin Neurophys* **103**: 554-562.

Elshaug, AG, Moss, JR and Southcott, AM (2005). Implementation of a split-night protocol to improve efficiency in assessment and treatment of obstructive sleep apnoea. Intern Med J **35**(4): 251-254.

Engleman, HM and Wild, M (2003). Improving CPAP use by patients with the sleep apnoea/hypopnoea syndrome (SAHS). Sleep Med Rev **7**(1): 81-99.

Farre, R, Montserrat, JM, Rigau, J, Trepat, X, Pinto, P and Navajas, D (2002). Response of automatic continuous positive airway pressure devices to different sleep breathing patterns: A bench study. Am J Respir Crit Care Med **166**: 469-473.

Gay, P, Herold, D and Olson, E (2003). A randomized, double-blind clinical trial comparing continuous positive airway pressure with a novel bilevel pressure system for treatment of obstructive sleep apnea syndrome. Sleep **26**(7): 864-869.

Grunstein, R (2005). Continuous positive airway pressure treatment for obstructive sleep apnea-hypopnea syndrome. *Principles and practice of sleep medicine (4th ed.)*. M. Kryger, T. Roth and W. Dement. Philadelphia, Elsevier Saunders: 1066-1080.

Harris, R (2003). Obstructive sleep apnoea syndrome: Symptoms and risk factors among Maori and non-Maori adults in Aotearoa. Masters of Public Health Thesis. Department of Public Health.University of Otago at the Wellington School of Medicine and Health Sciences. Wellington, New Zealand.

Johns, MW (1991). A new method for measuring daytime sleepiness: The Epworth Sleepiness Scale. Sleep **14**: 540-545.

Kazis, L, Anderson, J and Meenan, R (1989). Effect sizes for interpreting changes in health status. Med Care **27**(3 suppl): S178-S189.

Kribbs, N, Pack, AI, Kline, L, Smith, P, Schwartz, AR, Schubert, N, Redline, S, Henry, J, Getsy, J and Dinges, DF (1993). Objective measurement of patterns of nasal CPAP use by patients with obstructive sleep apnea. Am J Respir Crit Care Med **147**: 887-895.

Littner, M, Hirshkowitz, M, Davilla, D, Anderson, W, Kushida, C, Woodson, B, Johnson, S and Wise, M (2002). Practice parameters for the use of auto-titrating continuous positive airway pressure devices for titrating pressures and treating adult patients with obstructive sleep apnea syndrome: An American Academy of Sleep Medicine report. Sleep **25**(2): 143-147.

Massie, C, Hart, R, Peralez, K and Richards, G (1999). Effects of humidification on nasal symptoms and compliance in sleep apnea patients using continuous positive airway pressure.[comment]. Chest **116**: 403-408.

McArdle, N, Grove, A, Devereux, G, Mackay-Brown, L, Mackay, T and Douglas, N (2000). Split-night versus full-night studies for sleep apnoea/hypopnoea syndrome. Eur Respir J **15**(4): 670-675.

Mitler, MM, Gujavarty, KS and Browman, CP (1982). Maintenance of Wakefulness Test: A polysomnographic technique for evaluating treatment efficacy in patients with excessive somnolence. Electroenceph Clin Neurophys **53**(6): 658-661.

Neill, A, Wai, H, Bannan, S, Beasley, C, Weatherall, M and Campbell, A (2003). Humidified nasal continuous positive airway pressure in obstructive sleep apnoea. Eur Respir J **22**(2): 258-262.

Noseda, A, Kempenaers, C, Kerkhofs, M, Braun, S, Linkowski, P and Jann, E (2004). Constant vs auto-continuous positive airway pressure in patients with sleep apnea hypopnea syndrome and a high variability in pressure requirement. Chest **126**(1): 31-37.

Rechtschaffen, A and Kales, A (1968). A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Los Angeles, Brain information service/ Brain research Institute University of California.

Reeves-Hoche, M, Hudgel, D, Meck, R, Witteman, R, Ross, A and Zwillich, C (1995). Continuous versus bilevel positive airway pressure for obstructive sleep apnea. Am J Respir Crit Care Med **151**(2): 443-449.

Resta, O, Guido, P, Picca, V, Sabato, R, Rizzi, M, Scarpelli, F and Sergi, M (1998). Prescription of nCPAP and nBiPAP in obstructive sleep apnoea syndrome: Italian experience in 105 subjects. A prospective two centre study. Resp Med **92**: 820-827.

Rodway, G and Sanders, M (2003). The efficacy of split-night sleep studies. Sleep Med Rev **7**(5): 391-401.

Sanders, M, Costantino, J, Strollo, P, Studnicki, K and Atwood, C, Jr (2000). The impact of split-night polysomnography for diagnosis and positive pressure therapy titration on treatment acceptance and adherence in sleep apnea/hypopnea. Sleep **23**(1): 17-24.

Schafer, H, Ewig, S, Hasper, E and Luderitz, B (1998). Failure of CPAP therapy in obstructive sleep apnoea syndrome: Predictive factors and treatment with bilevel-positive airway pressure. Resp Med **92**: 208-215.

Schultz, K and Grimes, D (2002). Generation of allocation sequences in randomised trials: Chance, not choice. Lancet **359**: 515-519.

Sullivan, CE, Issa, FG, Bethon-Jones, M and Eves, L (1981). Reversal of obstructive sleep apnoea by continuous positive airway pressure applied through the nares. Lancet **1**: 862-865.

Wiest, G, Fuchs, F, Brueckl, W, Nusko, G, Harsch, I, Hahn, E and Ficker, J (2000). In vivo efficacy of heated and non-heated humidifiers during nasal continuous positive airway pressure (nCPAP)-therapy for obstructive sleep apnoea. Respir Med **94**(4): 364-368.

Yamashiro, Y and Kryger, M (1995). CPAP titration for sleep apnea using a split-night protocol. Chest **107**(1): 62-66.