Comparison of therapeutic and subtherapeutic nasal continuous positive airway pressure for obstructive sleep apnoea: a randomised prospective parallel trial

Crispin Jenkinson, Robert J O Davies, Rebecca Mullins, John R Stradling

Summary

Background Nasal continuous positive airway pressure (NCPAP) is widely used as a treatment for obstructive sleep apnoea. However, to date there are no randomised controlled trials of this therapy against a well-matched control. We undertook a randomised prospective parallel trial of therapeutic NCPAP for obstructive sleep apnoea compared with a control group on subtherapeutic NCPAP.

Methods Men with obstructive sleep apnoea, defined as an Epworth sleepiness score of 10 or more and ten or more dips per h of more than 4% SaO_2 caused by obstructive sleep apnoea on overnight sleep study, were randomly assigned therapeutic NCPAP or subtherapeutic NCPAP (about 1 cm H_2O) for 1 month. Primary outcomes were subjective sleepiness (Epworth sleepiness score), objective sleepiness (maintenance of wakefulness test), and SF-36 questionnaire measurements of self-reported functioning and well-being.

Findings 107 men entered the study: 53 received subtherapeutic NCPAP and 54 therapeutic NCPAP. Use of NCPAP by the two treatment groups was similar: 5·4 h (therapeutic) and 4·6 h (subtherapeutic) per night. Subtherapeutic NCPAP did not alter the overnight number of SaO_2 dips per h compared with baseline, and thus acted as a control. Therapeutic NCPAP was superior to subtherapeutic NCPAP in all primary outcome measures. The Epworth score was decreased from a median of 15·5 to 7·0 on therapeutic NCPAP, and from 15·0 to 13·0 on subtherapeutic NCPAP (between treatments, $p<0·0001$). Mean maintenance-of-wakefulness time increased from 22·5 to 32·9 min on therapeutic NCPAP and, not significantly, from 20·0 to 23·5 min on subtherapeutic NCPAP (between treatments $p<0·005$). Effect sizes for SF-36 measures of energy and vitality were 1·68 (therapeutic) and 0·97 (subtherapeutic) NCPAP (between treatments $p<0·0001$). For mental summary score, the corresponding values were 1·02 and 0·4 (between treatments $p=0·002$).

Interpretation Therapeutic NCPAP reduces excessive daytime sleepiness and improves self-reported health status compared with a subtherapeutic control. Compared with controls, the effects of therapeutic NCPAP are large and confirm previous uncontrolled clinical observations and the results of controlled trials that used an oral placebo.

Lancet 1999; **353:** 2100–05
See Commentary page ???

Health Services Research Unit, Division of Public Health and Primary Health Care, University of Oxford Institute of Health Sciences, Oxford (C Jenkinson DPhil)**; and Osler Chest Unit, Churchill Campus, Oxford Radcliffe Trust, Oxford, OX3 7LJ, UK** (R J O Davies MRCP, R Mullins RGN, J R Stradling FRCP)

Correspondence to: Dr John R Stradling
(e-mail: jstrad@resin.ox.ac.uk)

Introduction

Obstructive sleep apnoea is caused by airway occlusion during sleep, secondary to pharyngeal collapse. Each episode of collapse is usually stopped by transient arousal from sleep, which is necessary to restore pharyngeal muscle tone and thus reopen the airway. Severely affected patients have hundreds of obstructive episodes, and thus hundreds of episodes of brief arousal and sleep disturbance, which lead to excessive daytime sleepiness. Obstructive sleep apnoea is about 15 times more common in men than in women, and was first described accurately in 1966.[1] At first, the only effective treatment was tracheostomy, although weight loss in obese patients is sometimes beneficial. Sullivan and colleagues[2] showed that splinting open the upper airway during sleep with nasal continuous positive airway pressure (NCPAP, about 10 cm H_2O) via a nasal mask prevented recurrent pharyngeal collapse, decreased sleep fragmentation, and improved sleep quality.[3] Uncontrolled case series have suggested benefit of NCPAP, although none was a robust randomised controlled trial.[4] Nonetheless, the number of patients with obstructive sleep apnoea treated with NCPAP has increased rapidly. An evidence-based review of NCPAP therapy[4] concluded, in the absence of adequate trial data, that there was no good evidence to support the use of NCPAP for obstructive sleep apnoea. As a consequence of this review, and despite other more favourable reviews,[5–7] funding in the UK was withdrawn or severely limited for the investigation and treatment of obstructive sleep apnoea.

The only trial data with adequate blinding on the use of NCPAP for obstructive sleep apnoea at the time of the review used a tablet as placebo in the control group.[8] That study was judged unsatisfactory since a tablet does not control adequately for the placebo effect of a physical therapy that involves wearing of a mask over the nose at night. We undertook a controlled trial to compare conventional NCPAP (therapeutic) with NCPAP at low pressure (subtherapeutic). We aimed to assess whether NCPAP reduces excessive daytime sleepiness and improves self-reported functioning and well-being, compared with an appropriate control identical to therapeutic NCPAP but with no clinically significant effect on nocturnal obstructive events.

Methods

Patients

The Oxford Sleep Unit takes patients referred with possible obstructive sleep apnoea from the South Midlands, UK: a third of patients are from the Oxford area. Referrals are made by general practitioners (36%), ear, nose, and throat surgeons (41%), or other hospital consultants (23%). Patients were eligible for our trial if they had excessive daytime sleepiness, sleep apnoea, and were men aged between 30 and 75 years. Excessive daytime sleepiness was defined as an Epworth sleepiness score of 10 or more.[9] Obstructive sleep apnoea was defined as more than ten episodes per h of a greater than 4% fall in arterial oxygen saturation (SaO_2) during a sleep study, with confirmatory evidence

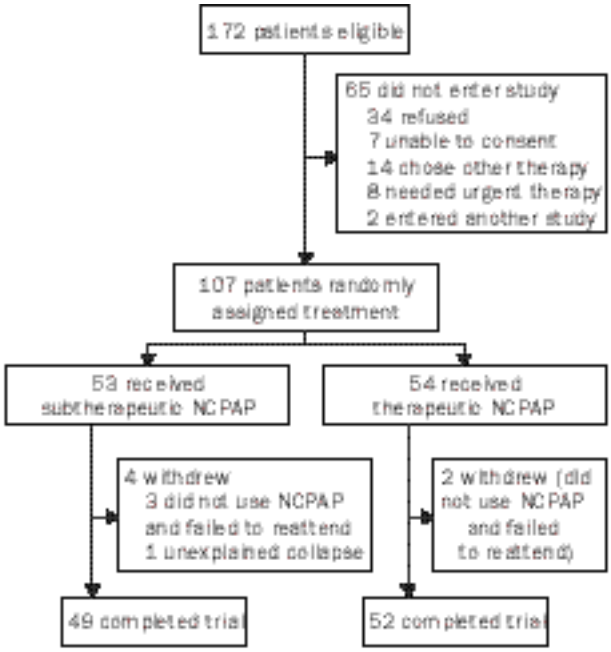
that these episodes were caused by pharyngeal collapse. All eligible patients entered the study unless they chose an alternative therapy (eg, weight loss, tonsillectomy), needed urgent NCPAP because of associated respiratory failure or because they were about to lose their job due to sleepiness, declined to participate, or had a mental disability that made informed consent impossible to obtain.

Eligibility for trial

Excessive subjective daytime sleepiness is the main indication for treatment of obstructive sleep apnoea in the UK. The Epworth sleepiness score,[9] which is the most widely used index to measure sleep apnoea subjectively, uses eight questions about the tendency to fall asleep in situations of differing stimulation (eg, watching television, talking to someone). Each question is scored from 0 to 3, to show increased tendency to fall asleep in each situation; the total score ranges from 0 (no sleepiness) to 24 (extremely sleepy): a score of 9 is the upper limit of normal.[10] Uncontrolled studies have shown that scores improve significantly after NCPAP treatment.[11]

We measured excessive daytime sleepiness objectively by use of a modified maintenance-of-wakefulness test.[12] During this test, patients are asked to resist sleep while semi-recumbent in a darkened room for up to 40 min, on four occasions in 1 day (0900 h, 1100 h, 1300 h, 1500 h). Patients are asked to stay awake but not to use active methods of keeping awake such as singing, shouting, or pinching themselves. Patients repeatedly tap a detector in response to a dim red light that flashes every 3 s. A computer controls the light and logs responses. Sleep is defined as failure to respond for 21 s (seven responses). After sleep onset the patient is awakened immediately. Mean ability to resist sleep onset is defined as mean time to sleep onset over the four periods, and was one of our primary endpoints. This test gives results similar to those of the conventional maintenance-of-wakefulness test, in which sleep onset is verified by encephalography rather than behavioural criteria.[13]

We assessed self-reported health status by use of the SF-36 questionnaire.[14] In a preparatory longitudinal cohort study of several health-status measures[15] the SF-36 had the best reliability, validity, and responsiveness for patients with obstructive sleep apnoea. SF-36 is a 36-item questionnaire that measures physical functioning, physical problems, emotional problems, social functioning, mental health, energy and vitality, pain, and general perception of health. For each variable, scores are coded, summed, and transformed onto a scale from 0 (worst possible health) to 100 (best possible health). Two summary scores are calculated—physical (physical component summary), and emotional well-being (mental component summary). These summary scores replicate the results from the original eight variables of the SF-36,[16] and are standardised such that a mean score of 50 (SD 10) reflects the mean score of the relevant population (our source was the Oxford Healthy Lifestyle Survey).[17] The SF-36 questionnaire has been used to measure decreased quality of life in several disorders including obstructive sleep apnoea.[15] On the basis of an earlier study,[15] we used the energy and vitality score and mental component summary as two further primary outcome measures in our study. Changes in each variable and the two component summaries were calculated by use of the effect size.[18] An effect size of 0·2 is considered small, 0·5 medium, and 0·8 large.[19] Since the SF-36 is a generic tool to assess self-reported health status, health gain from therapeutic interventions can be compared in different disorders.



Trial profile

We established obstructive sleep apnoea by a one-night sleep study that recorded patients' body movement and heart rate as markers of sleep disturbance, with arterial oxygen saturation measurements (SaO_2) and snoring as markers of respiratory impairment (Visi-Lab monitoring system, Stowood Scientific Systems, Oxford, UK).[20] Resting SaO_2 was measured after the patient had been sitting for 20 min—the time taken to answer the study questionnaire. A video recording of the whole night was made to confirm that abnormalities on the tracings were associated with obstructive sleep apnoea. Severity of obstructive sleep apnoea was measured by the number of falls in SaO_2 of more than 4% in each h of study. This measurement predicts the severity of obstructive sleep apnoea and its response to treatment as well as any other index.[20]

NCPAP trial

After the above tests, and after consent was obtained, eligible patients were admitted to hospital for a second night. All were shown a video about NCPAP. A specialist nurse taught each patient how to use the nasal masks, after which patients were randomly assigned either therapeutic or subtherapeutic NCPAP by use of a series of opaque sealed envelopes prepared in advance of the trial.

NCPAP was computer-controlled by use of a DeVilbiss Horizon autotitrating NCPAP machine (Sunrise Medical, Somerset, PA, USA), which uses algorithms based on snoring and reductions in breathing and which is as accurate as a skilled technician.[21] Subtherapeutic (control) NCPAP was identical to therapeutic NCPAP, except that pressure at the mask was unlikely

	Subtherapeutic NCPAP (n=49)	Therapeutic NCPAP (n=52)	p
Age (years)	48 (36–68)	50 (33–71)	0·84
Weight (kg)	109 (80·5–160)	105 (80·2–144)	0·32
Body-mass index (kg/m^2)	35·0 (26·9–51·4)	35·1 (25·8–44·3)	0·23
Neck circumference (cm)	45·7 (41·4–52·1)	44·5 (38·9–51·3)	0·18
>4% SaO_2 (dips/h)	28·5 (10·7–68·7)	32·9 (15·5–63·4)	0·30
Epworth sleepiness score	17·0 (10·0–23·0)	16·0 (10·7–21·7)	0·81
Maintenance of wakefulness test (min)	20·0 (3·5–40·0)	22·5 (7·6–40·0)	0·49
SF-36 mental component summary*	43·5 (10·7)	44·8 (10·4)	0·48
SF-36 physical component summary*	42·6 (10·1)	43·7 (11·6)	0·43
Daytime SaO_2 (%)	95·0 (91·5–96·5)	95·0 (91·7–97·0)	0·49
Final CPAP pressure (cm H_2O)	9·3 (6·0–15·0)	9·0 (6·8–13·4)	0·59

Date are median (5th–95th centiles) except * mean (SD).

Table 1: **Characteristics of patients**

	Subtherapeutic NCPAP			Therapeutic NCPAP			p for difference between endpoints after therapy	Difference		p for difference between differences
	Before	After	p before/after	Before	After	p before/after		Subtherapeutic	Therapeutic	
Epworth sleepiness score	15·0 (9·0–22·5)	13·0 (4·0–19·0)	<0·0001	15·5 (10·0–23·0)	7·0 (0·7–17·0)	<0·0001	<0·0001	−2·0 (−12·0 to 4·0)	−9·0 (−19·0 to 1·4)	<0·0001
Maintenance of wakefulness test	20·0 (3·5–40·0)	23·5 (7·0–40·0)	ns	22·5 (7·6–40·0)	32·9 (11·6–40·0)	<0·0001	0·002	0 (−14·3 to 23·6)	6·75 (−16·4 to 25·6)	0·005
Daytime % SaO_2*	95·0 (91·5–96·5)	95·0 (92·0–96·5)	ns	95·0 (91·7–97·0)	96·0 (93·7–97·4)	<0·0001	0·001	0 (−2·0 to 2·0)	1·0 (−1·0 to 3·4)	0·020
NCPAP use (h)	..	4·6 (0·7–8·5)	..	..	5·4 (2·2–7·4)	..	..	..	..	0·035

Data are median (5th–95th centiles). *% arterial oxygen saturation by oximetry. ns=not significant.

Table 2: **Sleepiness, awake % SaO_2, and NCPAP use before and after therapeutic and subtherapeutic NCPAP**

to be enough to splint open the pharynx. Subtherapeutic pressure (about 1 cm H_2O) was achieved by use of a NCPAP machine set to the lowest pressure possible, and by partly restricting airflow within the machine. Six extra 4 mm holes were cut in the rubber collar of the connecting tube at the mask end, to allow more air to escape and keep the nasal pressure low while ensuring no CO_2 reinhalation.

The morning after the NCPAP trial, patients were sent home with a therapeutic or subtherapeutic NCPAP machine, according to the randomisation schedule. All patients had telephone access to specialist nurses for advice if required. At 4 weeks, the patients were readmitted for a repeat of the daytime assessments of sleepiness and self-reported health status. The time clocks on the NCPAP machines were read to calculate mean nightly use over the month.

Patients were told that we were comparing two NCPAP pressures to find out which was the more effective in controlling their symptoms, and that one might be more effective than the other. Since they had never experienced NCPAP before, there was no reason for patients to realise that the lower pressure might be subtherapeutic. Therefore, combined with the parallel design, it was extremely unlikely that the patients would behave differently towards the two therapies in a way that would invalidate the masking. The research nurse (RM) who did the maintenance-of-wakefulness tests, the quality-of-life assessments, and the Epworth score was unaware of which NCPAP pressure each patient had received, and was not involved in provision of NCPAP management of patients, telephone support, or equipment maintenance. Our study was therefore effectively double-blinded, despite the physical nature of the therapy.

The Central Oxford Research Ethics Committee (number 96.127) approved the study protocol.

Statistical analysis

Our four primary outcomes were objective (maintenance-of-wakefulness) and subjective (Epworth) sleepiness, energy and vitality, and the mental component summary of the SF-36 questionnaire. Secondary endpoints included the seven other variables in the SF-36, the physical component summary, and awake resting arterial oxygen saturation. Previous uncontrolled studies that used the SF-36 to assess response to NCPAP[15] allowed a preliminary power calculation, which suggested that about 150 patients would need to be randomly assigned treatment. A planned interim analysis when 40 patients had been randomised[22] suggested that a sample size of 100 should show a difference between subtherapeutic and therapeutic NCPAP in all four primary outcome measures at significance of at least p=0·01. The interim analysis was done primarily to ensure that symptoms of control patients were not being made worse by subtherapeutic NCPAP pressure.

All data were computer-analysed with SPSS (version 7.5.1). Data are expressed as median (5th to 95th centiles) because all measures have upper and lower limits to their values and are not normally distributed, except for SF-36 data, which are given as mean (SD). Changes with treatment were assessed in three ways. Differences between before and after therapy were compared in each group by Wilcoxon matched-pairs signed-ranks tests. The change in outcome measures after treatment were compared between the groups by Wilcoxon rank-sum test. Actual outcome measures in the two groups after treatment were compared by Wilcoxon rank-sum test, this being the most rigorous test of any differences. Significance for the four primary endpoints was set at 0·01. The relation between changes in Epworth sleepiness score after treatment and actual nightly use of NCPAP was assessed by Pearson's correlation coefficient.

Results

During the recruitment period (January, 1997, to August, 1998) 172 patients were eligible for the study (figure 1). 65 people were excluded: 34 refused, mostly because of the longer study time or long-distance travel. Seven were judged too mentally impaired to give informed consent (three with major psychoses, two with severe learning difficulties, two with alcohol dependence), 14 chose alternative therapies or believed that their symptoms did not warrant such an intrusive therapy; eight needed urgent therapy (respiratory failure, imminent job loss due to sleepiness and inability to drive); and two entered a different study. Six patients who refused to return for the 1-month follow up (four subtherapeutic, two therapeutic) were not analysed.

Baseline characteristics of the two groups were similar (table 1). In a sample of 26 trial patients on subtherapeutic

	Subtherapeutic NCPAP			Therapeutic NCPAP			p for difference between endpoints after therapy	Effect size		p comparing effect sizes
	Before	After	p before/after	Before	After	p before/after		Subtherapeutic	Therapeutic	
General health perception	59·5 (20·4)	62·2 (22·8)	ns	59·2 (18·4)	70·5 (22·5)	<0·0001	ns	0·13	0·61	0·002
Physical functioning	78·6 (22·1)	80·1 (21·7)	ns	80·9 (22·7)	85·9 (23·1)	0·005	ns	0·07	0·22	ns
Social functioning	73·0 (26·1)	82·3 (23·6)	ns	73·5 (26·1)	91·2 (18·3)	0·0001	ns	0·36	0·68	ns
Physical role	58·7 (37·0)	70·9 (38·0)	ns	62·0 (37·2)	90·9 (23·2)	<0·0001	0·002	0·33	0·78	ns
Mental role	68·7 (36·3)	73·5 (35·3)	ns	73·7 (33·2)	93·6 (19·9)	0·0004	0·0002	0·13	0·60	ns
Bodily pain	76·2 (25·5)	83·9 (23·4)	ns	82·1 (23·8)	91·2 (13·4)	0·008	ns	0·30	0·39	ns
Mental health	68·7 (18·2)	75·8 (18·0)	0·01	73·2 (16·8)	85·5 (12·7)	<0·0001	0·002	0·39	0·73	ns
Energy and vitality	33·9 (17·5)	50·9 (20·5)	<0·0001	35·4 (22·4)	73·0 (17·0)	<0·0001	<0·0001	0·97	1·68	<0·0001
Mental component summary	43·5 (10·7)	47·8 (10·1)	0·01	44·8 (10·4)	55·4 (7·0)	<0·0001	5×10^{-5}	0·40	1·02	0·002
Physical component summary	42·6 (10·1)	45·5 (10·4)	0·007	43·7 (11·6)	49·4 (10·1)	<0·0001	0·009	0·29	0·49	0·080

Data are mean (SD). ns=not significant.

Table 3: **SF-36 scores before and after therapeutic and subtherapeutic NCPAP**

NCPAP, mean number of falls of more than 4% SaO_2 per h was 33 (5–95% CI 6–72) compared with 29 (11–69) on the diagnostic night (difference not significant, and not different from the group as a whole, table 1). In addition, the Epworth score and number of 4% falls below SaO_2 per h of the 65 people not randomised were no different from the 101 patients. We included the final NCPAP pressures required to control obstructive sleep apnoea on the second titration study, after the trial had finished, as an alternative index of severity of obstructive sleep apnoea, to give a quantitative measure of the tendency for the pharynx to collapse.

Table 2 shows the data for objective and subjective sleepiness before and after subtherapeutic and therapeutic NCPAP. There was a small but significant effect of subtherapeutic NCPAP on the subjective measure of sleepiness (Epworth) only. Therapeutic NCPAP gave large improvements in both measures of sleepiness that were significantly greater than subtherapeutic NCPAP, whichever of the analyses was used. The use of NCPAP by the subtherapeutic group was 48 min per night less than that of the therapeutic group (p=0·035). The small improvement in daytime SaO_2 after therapeutic NCPAP was significant compared with subtherapeutic NCPAP (p=0·02).

Table 3 shows data for the eight SF-36 variables and the two component summaries (mental, physical). The energy and vitality variable showed the largest difference between therapeutic and subtherapeutic NCPAP. General health perception showed the second-largest difference. As with the Epworth score, there were small but significant effects on the energy and vitality variable within the control group on subtherapeutic NCPAP.

Thus, for all primary outcome measures there was significant physiological and clinical benefit from therapeutic NCPAP compared with subtherapeutic NCPAP. There was also a correlation in the therapeutic NCPAP group between mean nightly use of NCPAP machines and improvement in Epworth score (r=0·60, p<0·0001), but not in the subtherapeutic group (r=0·15, p=0·3).

Discussion

Therapeutic NCPAP has a clear advantage over subtherapeutic NCPAP. Compared with a control, the main symptom of obstructive sleep apnoea (sleepiness) was significantly improved by therapeutic NCPAP, both objectively and subjectively. Therapeutic NCPAP improved Epworth scores by 7 points (95% CI 5–9), and for 73% of the therapeutic NCPAP group Epworth scores returned to within the normal range (number needed to treat=1·4), compared with only 29% of the subtherapeutic control group. These improvements correlated significantly with the amount of time that patients actually used their NCPAP machines, but only in the therapeutic group.

The improvement in median maintenance-of-wakefulness test time with therapeutic NCPAP was about 7 min greater than that with subtherapeutic NCPAP (95% CI 3–11). Before therapeutic NCPAP, only 8% could stay awake for the full 40 min of the test on all four occasions, whereas 37% could do so after treatment (subtherapeutic NCPAP 18% and 10%, respectively). The mean test time after therapeutic NCPAP was 30 min, which is similar to the mean in healthy people of about 35 min.[12] Thus, these patients approached normal scores for objectively measured sleep resistance after only 4 weeks' treatment. The 26 (50%) therapeutic NCPAP users with the best compliance (>5·4 h/night) had a median maintenance-of-wakefulness score after treatment of 38·3 min (mean 33·0 min), which is similar to that for healthy people and 12 min longer (95% CI 7·5–6·5) than that after subtherapeutic controlled NCPAP. A randomised controlled trial[23] to assess the use of modafinil, recently licensed in the UK for the treatment of hypersomnolence due to narcolepsy, showed that modafinil improved maintenance-of-wakefulness scores from 6 min to 9 min. Thus the therapeutic effect of NCPAP on objective sleepiness caused by obstructive sleep apnoea is substantially greater than the effect of modafinil on sleepiness caused by narcolepsy. Two randomised controlled trials of NCPAP for obstructive sleep apnoea of a similar severity to our study[24] used an oral placebo, and both those studies showed a similar improvement in Epworth scores to our study. However, improvement in the objective measure of sleepiness compared with the placebo was smaller, which may have been due to less use of NCPAP by the patients (3·3 h/night).

Our previous uncontrolled work on the use of health-status questionnaires in assessment of NCPAP therapy suggested that in a controlled study large gains in self-reported functioning and well-being were likely when assessed with the SF-36.[15] Our study confirmed these predictions: the mental component summary of the SF-36 improved with an effect size of 1·02, and the value for energy and vitality was 1·68, where 0·8 is judged a large change.[19] The actual post-treatment figure for the mental component summary of 55·4 rose to more than the mean (50) for a matched population.[17] The SF-36 has been used to measure health gain in other disorders, such as Parkinson's disease and heart failure, and the effect sizes have been substantially smaller.[25,26]

In the subtherapeutic NCPAP group, we expected improvement (placebo effect) at 1 month in SF-36 energy and vitality scores, and a small improvement in Epworth scores. Symptoms of the study population are chronic and can be disabling and depressing. In the clinic, patients are told that the diagnosis is now known and treatment available. They are admitted for a training session and shown a video that includes interviews with patients already on treatment. Thus, patients are likely to have high expectations of their treatment, which gives rise to a powerful placebo effect. Alternatively subtherapeutic pressure (1 cm H_2O) may have improved some aspect of obstructive sleep apnoea, although this explanation was not supported by unchanged results of home monitoring of overnight SaO_2 whilst on subtherapeutic NCPAP. If this latter explanation were true, then it would have lessened the experimental differences between therapeutic and subtherapeutic NCPAP and led to an underestimation of the benefit of NCPAP.

The improvement in daytime SaO_2 on therapeutic versus subtherapeutic NCPAP implies a physiological improvement in gas exchange, due either to an improvement in ventilation/perfusion matching in the lung or to increased ventilation. Without arterial blood gas estimations neither of these explanations can be eliminated, although previous uncontrolled work has shown improvements in $PaCO_2$ concentrations in patients with pretreatment hypercapnia after NCPAP treatment for obstructive sleep apnoea.[27] A change in SaO_2 from 95% to 96% represents a rise in PaO_2 of about 1 kPa.

A controlled trial of NCPAP for obstructive sleep apnoea[8] has been criticised for use of a tablet placebo,[4] since physical NCPAP therapy might have a more powerful effect than a tablet. For these reasons, our study used a control treatment as similar to normal NCPAP as possible, but without any measurable improvement in severity of obstructive sleep apnoea. Concerns have been expressed over the safety of subtherapeutic NCPAP,[28] but our design prevented patients from reinhalation of CO_2 and negative inspiratory pressures that would have loaded respiration further. There is a perceptible difference between our subtherapeutic pressure (1 cm H_2O) and the therapeutic pressure, so we used patients naïve to NCPAP and did not employ a crossover strategy. Adaptation to NCPAP is difficult for many patients, and any suspicion that they might have been given ineffective treatment would probably have lessened their enthusiasm to continue trying and thus have lessened any placebo effect. There was no reason for patients to suspect that they might have been on an ineffective pressure, although the slightly lower NCPAP usage by the subtherapeutic group might suggest less enthusiasm due to lack of efficacy. Furthermore, we ensured that the patients and the research nurse were unaware of treatment allocation, and the positive placebo response and only minimally lower NCPAP compliance are reassuring since they suggest that this masking was effective. We are therefore confident that this trial was as double-blind as is possible with a physical therapy.

Entry criteria for our study were inevitably a compromise. The correlation between conventional sleep-study measures of severity of obstructive sleep apnoea and sleepiness (or its subsequent improvement with NCPAP) is not good, and rarely exceeds 0·5,[20,29] probably because existing sleep-study indices of severity of obstructive sleep apnoea do not measure all aspects of sleep fragmentation and therefore cannot entirely predict subsequent effects on daytime sleepiness. There are many steps between obstructive respiratory events at night and a feeling of excessive sleepiness during the day. Events other than apnoeas can lead to waking (hypopnoeas, obstructive snoring), the degree of arousal required to clear the airway varies,[30] obstructive events may cluster or be spread out, extra sleep may be taken during the day, and differing lifestyles will alter the effects of sleep fragmentation on symptoms. Indices used to quantify obstructive sleep apnoea include decreased airflow at the nose and mouth, falls in SaO_2, and various markers of recurrent arousal from sleep. Despite theoretical differences between these indices, they are quite similar in their ability to predict excessive daytime sleepiness and its response to treatment with NCPAP[20] although, in general, measurements of SaO_2 dips per h are a little lower than equivalent apnoea/hypopnoea indices. So that our trial findings should be generalisable to UK practice, we identified patients for our study on the basis of the number of SaO_2 dips of more than 4% per h, since oximeters rather than full polysomnography are used most widely to quantify obstructive sleep apnoea in the UK. The Epworth questionnaire to quantify subjective sleepiness has been thoroughly validated and is extremely easy to use, so we decided on the simple criteria of subjective sleepiness above the normal range, and a severity marker for obstructive sleep apnoea based on oximetry. However, sleep units see patients with obstructive-sleep apnoea who have SaO_2 dip rates lower than the strict entry criteria for our study but who nonetheless show response to NCPAP through objective improvements in sleepiness. Thus, it would be wrong to use the results of our study to withhold treatment from patients who do not satisfy our trial entry criteria, but who do have proven recurrent upper-airway obstruction during sleep and disabling daytime sleepiness. More work is needed to define the lower end of the range of severity that is still likely to respond to treatment with NCPAP, although this work will be difficult owing to the complexities of measuring the severity of the disorder.

Contributors

Crispin Jenkinson and John Stradling were responsible for study proposal, design, execution, and analysis. Robert Davies was also responsible for study design and execution. Rebecca Mullins was responsible for entry of patients to the study, collection of clinical measurements, and data entry. John Stradling drafted the paper, and all investigators contributed to the final paper.

Acknowledgments

The study was supported by the NHS Executive (Anglia Oxford) grant number HSR/UOX/0296/72. We thank all the Oxford Sleep Unit staff for their assistance.

References

1 Guilleminault C, Tilkian A, Dement WC. The sleep apnea syndromes. *Ann Rev Med* 1976; **27:** 465–84.
2 Sullivan CE, Issa FG, Berthon-Jones M, Eves L. Reversal of obstructive sleep apnoea by continuous positive airway pressure applied through the nares. *Lancet* 1981; i: 862–65.
3 Issa FG, Sullivan CE. The immediate effects of nasal continuous positive airway pressure treatment on sleep pattern in patients with obstructive sleep apnoea syndrome. *Electroencephalogr Clin Neurophysiol* 1986; **63:** 10–17.
4 Wright J, Johns R, Watt I, Melville A, Sheldon T. Health effects of obstructive sleep apnoea and the effectiveness of continuous positive airways pressure: a systematic review of the research evidence. *BMJ* 1997; **314:** 851–60.
5 The Australian Health Technology Advisory Committee. The effectiveness and cost-effectiveness of nasal continuous positive airway pressure (nCPAP) in treatment of obstructive sleep apnea in adults. Canberra: National Health and Medical Research Council and the New Zealand Ministry of Health, 1996.
6 Fraser AK. Obstructive sleep apnoea and allied disorders. Glasgow: Scottish Needs Assessment Programme, Scottish Forum for Public Health Medicine, 1997.
7 The Wessex Institute for Health Research and Development. Continuous positive airway pressure for obstructive sleep apnoea. Southampton: University of Southampton, 1994.
8 Engleman HM, Martin SE, Deary IJ, Douglas NJ. Effect of continuous positive airway pressure treatment on daytime function in sleep apnoea/hypopnoea syndrome. *Lancet* 1994; **343:** 572–75.
9 Johns MW. Daytime sleepiness, snoring, and obstructive sleep apnea: the Epworth sleepiness scale. *Chest* 1993; **103:** 30–36.
10 Johns M, Hocking B. Daytime sleepiness and sleep habits of Australian workers. *Sleep* 1997; **20:** 844–49.
11 Hardinge FM, Pitson DJ, Stradling JR. Use of the Epworth Sleepiness Scale to demonstrate response to treatment with nasal continuous positive airway pressure in patients with obstructive sleep apnoea. *Respir Med J* 1995; **89:** 617–20.
12 Doghramji K, Mitler MM, Sangal RB, et al. A normative study of the maintenance of wakefulness test (MWT). *Electroencephalogr Clin Neurophysiol* 1997; **103:** 554–62.
13 Bennett LS, Stradling JR, Davies RJO. A behavioural test to assess daytime sleepiness in obstructive sleep apnoea. *J Sleep Res* 1997; **6:** 142–45.
14 Tarlov AR, Ware JE Jr, Greenfield S, Nelson EC, Perrin E, Zubkoff M. The Medical Outcomes Study: an application of methods for monitoring the results of medical care. *JAMA* 1989; **262:** 925–30.
15 Jenkinson C, Stradling J. Petersen S. Comparison of three measures of quality of life outcome in the evaluation of continuous positive airways pressure therapy for sleep apnoea. *J Sleep Res* 1997; **6:** 199–204.
16 Ware JE Jr, Kosinski M, Bayliss MS, McHorney CA, Rogers WH, Raczek A. Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: summary of results from the Medical Outcomes Study. *Med Care* 1995; **33:** AS264–79.
17 Wright L, Harwood D, Coulter A. Health and lifestyles in the Oxford Region. Oxford: Health Services Research Unit, University of Oxford, 1992.
18 Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989; **27:** S178–89.

19 Cohen J. Statistical power for the behavioral sciences. New York: Academic Press, 1977.
20 Bennett LS, Langford BA, Stradling JR, Davies RJO. Sleep fragmentation indices as predictors of daytime sleepiness and nCPAPA responce in OSA. *Am J Respir Crit Care Med* 1998; **158:** 778–86.
21 Stradling JR, Barbour C, Pitson DJ, Davies RJO. Automatic nasal continuous positive airway pressure titration in the laboratory: patient outcomes. *Thorax* 1997; **52:** 72–75.
22 Stradling JR, Jenkinson C, Davies RJ, Mullins B. Randomised, sham-placebo controlled, parallel study of nasal continuous positive airway pressure (NCPAP) for the treatment of obstructive sleep apnoea (OSA)—interim analysis. *Am J Respir Crit Care Med* 1998; **157:** A344 (abstr).
23 US Modafinil in Narcolepsy Multicentre Study Group. Randomised trial of modafinil for the treatment of pathological somnolence in narcolepsy. *Ann Neurol* 1998; **43:** 88–97.
24 Douglas NJ. Systematic review of the efficacy of nasal CPAP. *Thorax* 1998; **53:** 414–15.
25 Jenkinson C, Peto V, Fitzpatrick R, Greenhall R, Hyman N. Self-reported functioning and well-being in patients with Parkinson's disease: comparison of the short-term health survey (SF-36) and the Parkinson's Disease Questionnaire (PDQ-39). *Age Ageing* 1995; **24:** 505–09.
26 Jenkinson C, Jenkinson D, Shepperd S, Layte R, Petersen S. Evaluation of treatment for congestive heart failure in patients aged 60 years and older using generic measures of health status (SF36 and COOP charts). *Age Ageing* 1997; **26:** 7–13.
27 Aubert-Tulkens G, Willems B, Veriter CL, Coche E, Stanescu DC. Increase in ventilatory response to CO_2 following tracheostomy in obstructive sleep apnea. *Bull Eur Physiopathol Respir* 1980; **16:** 587–93.
28 Krieger J, Weitzenblum E, Monassier JP, Stoeckel C, Kurtz D. Dangerous hypoxaemia during continuous positive airway pressure treatment of obstructive sleep apnoea. *Lancet* 1983; ii: 1429–30.
29 Cheshire K, Engleman H, Deary I, Shapiro C, Douglas NJ. Factors impairing daytime performance in patients with the sleep apnoea/hypopnoea syndrome. *Arch Intern Med* 1992; **152:** 538–41.
30 Stradling JR, Pitson DJ, Bennett LS, Barbour C, Davies RJO. Variation in the arousal pattern following obstructive events in obstructive sleep apnoea. *Am J Respir Crit Care Med* 1999; **159:** 130–36.

Prediction of benefit from carotid endar terectom y in individual patients: a risk-modelling study

P M Rothwell, C P Warlow, on behalf of the European Carotid Surgery Trialists' Collaborative Group

Summary

Background Carotid endarterectomy lowers the risk of carotid territory ipsilateral ischaemic stroke, and is the treatment of choice, in patients with recently symptomatic 70–99% carotid stenosis. However, the 3-year risk of stroke on medical treatment alone is only about 20%. We investigated whether the efficacy of endarterectomy would be improved if patients with a high risk of stroke on medical treatment and a low risk of operative stroke or death could be identified.

Methods We developed two prognostic models from data on patients with 0–69% carotid stenosis in the European Carotid Surgery Trial (ECST). The medical model predicted risk of ipsilateral carotid territory major ischaemic stroke (fatal or lasting longer than 7 days) on medical treatment and the surgical model predicted risk of major stroke and death within 30 days of endarterectomy. From these models we developed a prognostic score to identify patients with a high risk of stroke on medical treatment but a low operative risk. We validated the models and tested the scoring system on 990 ECST patients with 70–99% carotid stenosis assigned surgery (594) or medical treatment only (396).

Findings When patients with 70–99% stenosis were stratified by the scoring system, which was based on seven independent prognostic factors, endarterectomy was beneficial in only 162 (16%) patients. The 5-year absolute risk of carotid territory ipsilateral major ischaemic stroke, operative major stroke, or death was lowered by 33% in the 16% of patients with a score of 4 or more (odds ratio 0·12 [95% CI 0·05–0·29], p<0·0001), but not in the other 828 (84%) patients (1·00 [0·65–1·54], p=0·7).

Interpretation Many patients with recently symptomatic 70–99% carotid stenosis may not benefit from carotid endarterectomy. Validation of the predictive score is needed on external datasets, but risk-factor modelling could be useful to identify those patients in whom endarterectomy will be beneficial.

Lancet 1999; **353:** 2105–10

Department of Clinical Neurology, Radcliffe Infirmary, Oxford OX2 6HE, UK (P M Rothwell MD)**; and Department of Clinical Neurosciences, Western General Hospital, Edinburgh** (Prof C P Warlow FRCP)

Correspondence to: Dr P M Rothwell (e-mail: peter.rothwell@clneuro.ox.ac.uk)

Introduction

Atherothrombotic stenosis at or around the carotid bifurcation is associated with an increased risk of ipsilateral carotid territory ischaemic stroke. This risk is lowered in some patients by carotid endarterectomy.[1–5] Two randomised controlled trials[2,3] of carotid endarterectomy with best medical treatment versus best treatment alone in patients with recently symptomatic carotid stenosis have produced clear results. The European Carotid Surgery Trial (ECST) and the North American Symptomatic Carotid Endarterectomy Trial (NASCET) showed clear decreases in the overall risk of stroke in operated patients with recently symptomatic severe (70–99%) carotid stenosis.[2,3] ECST also showed that surgery is harmful in patients with less than 70% stenosis.[4] NASCET showed no benefit in patients with 30–49% stenosis, but did show a small benefit in those with 50–69% stenosis.[5] However, the trials used different methods to measure the degree of stenosis on angiograms—the NASCET method underestimated stenosis compared with the ECST method.[6] The 50–69% stenosis group in which NASCET reported some benefit from surgery is equivalent to the 70–80% stenosis group in ECST.[7]

These trials therefore showed an overall benefit from endarterectomy in patients with recently symptomatic stenosis of 70–99% by the ECST method. However, this finding is of little help to the clinician who has to make decisions for individual patients. Although endarterectomy lowers the overall risk of ischaemic stroke by about 50% over the next 3 years in patients with 70–99% stenosis, only about 20% of such patients have a major stroke on medical treatment alone.[2,3] Therefore, surgery is of no value in 80%