

# Mandibular advancement appliance for obstructive sleep apnoea: results of a randomised placebo controlled trial using parallel group design

NIELS PETRI<sup>1,4</sup>, PALLE SVANHOLT<sup>2</sup>, BENI SOLOW<sup>2</sup>, GORDON WILDSCHIØDTZ<sup>3</sup> and PER WINKEL<sup>4</sup>

<sup>1</sup>ENT Department, Nykøbing F. Hospital, Region Zealand, Nykøbing F., <sup>2</sup>Department of Orthodontics, University of Copenhagen, Copenhagen, <sup>3</sup>Centre of Expertise, Oringe Psychiatric Hospital, Vordingborg and <sup>4</sup>Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

Accepted in revised form 15 January 2008; received 4 July 2007

**SUMMARY** The aim of this trial was to evaluate the efficacy of a mandibular advancement appliance (MAA) for obstructive sleep apnoea (OSA). Ninety-three patients with OSA and a mean apnoea–hypopnoea index (AHI) of 34.7 were centrally randomised into three, parallel groups: (a) MAA; (b) mandibular non-advancement appliance (MNA); and (c) no intervention. The appliances were custom made, in one piece. The MAAs had a mean protrusion of the mandible of 74% (range 64–85%). Outcome measures, assessed after continuous use for 4 weeks, were AHI (polysomnography), daytime sleepiness (Epworth) and quality of life (SF-36). Eighty-one patients (87%) completed the trial. The MAA group achieved mean AHI and Epworth scores significantly lower ( $P < 0.001$  and  $P < 0.05$ ) than the MNA group and the no-intervention group. No significant differences were found between the MNA group and the no-intervention group. The MAA group had a mean AHI reduction of 14.1 (95% CI 7.4–20.8), and a mean Epworth score reduction of 3.3 (95% CI 1.8–4.8). Eight MAA patients (30%) achieved a reduction in AHI  $\geq 75\%$  ending with an AHI  $< 5$ , half of them having baseline AHI  $> 30$ . Sensitivity analyses confirmed these results. MAA had a significant beneficial effect on the vitality domain of SF-36. Four MAA patients (14.8%) and two MNA patients (8%) discontinued interventions because of adverse effects. Our conclusion is that MAA has significant beneficial effects on OSA, including cure in some cases of severe OSA. Protrusion of the mandible is essential for the effect. MNA has no placebo effect. MAA may be a good alternative to CPAP in subsets of OSA patients.

**KEYWORDS** Epworth Sleepiness Scale, oral appliances, quality of life (SF-36), sleep apnoea

## INTRODUCTION

Obstructive sleep apnoea (OSA) is by far the most common sleep-related breathing disorder, affecting 2–4% of the adult population (Young *et al.*, 1993). The obstructions, located in

*Correspondence:* Dr Niels Petri, ENT Department, Nykøbing F. Hospital, Region Zealand, Fjordvej 15, DK-4800 Nykøbing F., Denmark. Tel.: +45 56 51 56 34; fax: +45 56 51 57 71; e-mail: npe@cnf.stam.dk and nipe@regionsjaelland.dk

the pharyngeal airway, are considered to lead to sleep fragmentation, resulting in excessive daytime sleepiness, which has consequences for the ability to work, for road safety and for quality of life. Furthermore, OSA is an independent risk factor for cardiovascular morbidity (Lavie *et al.*, 2000; Marin *et al.*, 2005; Peker *et al.*, 2002; Peppard *et al.*, 2000) and for mortality (Yaggi *et al.*, 2005). Effective treatment of OSA can reduce health service costs (Kapur *et al.*, 1999; Peker *et al.*, 1997).

The two current, dominant treatment modalities – continuous positive airway pressure (CPAP) and uvulopalatopharyngoplasty – were both introduced in 1981 (Fujita *et al.*, 1981; Sullivan *et al.*, 1981). Over time, enthusiasm for uvulopalatopharyngoplasty has diminished because of unsatisfactory results and lack of evidence supporting this surgery (Sundaram *et al.*, 2005). Today, the treatment of choice is CPAP. The main problem with CPAP, however, is unsatisfactory patient compliance (McArdle *et al.*, 1999; Popescu *et al.*, 2001; Waldhorn *et al.*, 1990). This problem has led to a search for alternative, more user-friendly treatment modalities.

Many orthodontists claim that oral appliances (OA) have good effect on OSA. The appliances are intended to enlarge the pharyngeal airway, directly by retaining the tongue, or indirectly by advancing the mandible. Another theory is that the appliances may cause stretch-induced activation of the pharyngeal motor system, thereby reducing soft tissue laxity and airway collapse (Talmant *et al.*, 1998; Tsuiki *et al.*, 2004). In recent years, several randomised trials were performed, most of them using CPAP alone or together with placebo tablet or conservative treatment as control (Barnes *et al.*, 2004; Engleman *et al.*, 2002; Ferguson *et al.*, 1996; Lam *et al.*, 2007; Tan *et al.*, 2002). Few randomised trials have used a sham OA as control (Hans *et al.*, 1997; Mehta *et al.*, 2001; Gotsopoulos *et al.*, 2002; Johnston *et al.*, 2002; Blanco *et al.*, 2005). These trials offer some evidence of a beneficial effect of OAs in patients with mild to moderate sleep apnoea, but they all have weaknesses, such as small sample size, crossover design, under-reporting of methods and data, and lack of blinding (Lim *et al.*, 2006).

The purpose of the present trial was to provide better evidence for the beneficial and harmful effects of OAs.

## METHODS

### Patients

The patients were consecutively recruited from a patient population of 483 patients referred for suspected OSA to the ENT clinic, a regional sleep centre at the Nykøbing F. Hospital. Inclusion criteria were: (1) an apnoea–hypopnoea index (AHI) over 5, found with a diagnostic polysomnography; (2) age more than 20 years; (3) sufficient set of teeth to hold a splint; and (4) written informed consent. Exclusion criteria were: (1) severe somatic or psychiatric disease; (2) periodontal disease; (3) temporomandibular dysfunction; and (4) pregnancy. All included patients had been offered CPAP treatment, but preferred OAs.

The patients underwent a standardised ENT examination by the clinical investigator (NP) with the patient in sitting position and an orthopantomography to establish that the teeth were healthy.

### Trial design

The trial had a three-armed, parallel-group design with central randomisation. The allocation was computer generated, taking

account of minimisation, stratification by sex and AHI above and below 30. Estimation of sample size was based on  $\alpha = 0.05$ ,  $\beta = 0.20$ , a minimal relevant difference of 15 on AHI and a standard deviation of 20, resulting in 29 patients per intervention group. Accordingly, it was decided to go for 30 patients per intervention group. When the patient was ready for inclusion at the Department of Orthodontics, University of Copenhagen, the Copenhagen Trial Unit was contacted via telephone for allocation to a mandibular advancement appliance (MAA), a mandibular non-advancement appliance (MNA) or no-intervention.

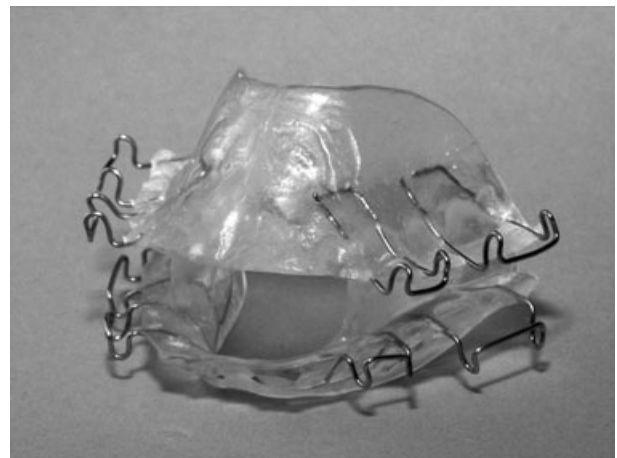
The protocol of the trial was approved by the regional research ethics committee (ID: 1998-1-34) and the Danish Data Protection Agency, and registered at ClinicalTrials.gov (ID: NCT00243139).

### Interventions

The appliances were one-piece, custom-made acrylic dental devices, the acrylic covering only the molars and premolars. The appliance was secured to these teeth by four stainless steel Adams clasps in each jaw (Fig. 1).

The acrylic plate was made as thin as possible and had a wide opening in the front to leave maximal space for the tongue. The MAA advanced the mandible to the most protrusive position without discomfort with a 5-mm vertical opening in front. The MNA was holding the mandible in the occlusal position.

Using a George Gauge™ instrument (George, 1992), measurements were taken for the: (1) most retrusive; (2) occlusal; (3) maximal protrusive position of the mandible; and (4) most protrusive position without discomfort. In case of discomfort, the protrusion and frontal opening was adjusted by rebuilding the appliance. The final mean protrusion was 74% (range 64–85%) of the difference between the maximal protrusive and retrusive positions. Because of inconvenience of the Adams clasps, in eight cases the appliances were rebuilt and secured to all the teeth by buccal acrylic extensions. A single dentist (PS) treated all the patients, and one dental technician made all the



**Figure 1.** The mandibular advancement appliance.

appliances. Whether a MAA or MNA was delivered, the patients as well as the clinical investigator were blinded. The no-intervention group was unblinded.

### Baseline and outcome measurements

The intervention effects were measured by polysomnography and questionnaires about daytime sleepiness, the Epworth Sleepiness Scale (Johns, 1991) and quality of life (QOL), the Short-Form General Health Survey, SF-36 (Ware, 1996). Clinical snoring disturbances were measured by a questionnaire about witnessed apnoeas, snoring frequency and snoring intensity, assessed on a four-point scale. The baseline measurements were performed just before the start of the intervention and the outcome measurements after 4 weeks in the no-intervention group. In the other two groups, the measurements were performed after 4 weeks of continuous use following the completion of final adjustments of fit.

The sleep studies were performed unattended in the patients' homes, using overnight polysomnography recorded by Embla A10 Recording System (Embla, Amsterdam, the Netherlands). This system continuously recorded electroencephalography (EEG), electrooculography, electromyography submentally, oronasal airflow (thermistors), thoraco-abdominal movements (piezo sensor), body position, snoring (vibrations on the neck) and finger pulse oximetry. Calculated respiratory events were AHI, supine and non-supine AHI, the percentage of apnoeas and hypopnoeas, the average duration of apnoeas and hypopnoeas, and the average oxygen desaturation. Apnoea was defined as cessation of airflow with persistent respiratory effort for at least 10 s. Hypopnoea was defined as clear amplitude reduction in measures of breathing for at least 10 s terminated by abrupt recovery breath and associated with either oxygen desaturation of  $\geq 3\%$  or arousal.

Sleep stages and respiratory events were scored manually by a single sleep investigator (GW), blinded to the group assignment, and using standard criteria for sleep stages (Rechtschaffen and Kales, 1968).

The primary outcome measure was the change in AHI. For comparison with other studies, we also calculated the percentage of patients achieving: (1) reduction in AHI  $\geq 50\%$ ; (2) AHI  $< 10$ ; and (3) AHI  $< 5$ , together with the success criteria according to Mehta *et al.*, (2001), defining complete response as resolution of symptoms and outcome AHI  $< 5$ , and partial response as improvement of symptoms and reduction in AHI  $\geq 50\%$  but outcome AHI  $\geq 5$ . Failure was defined as a reduction in AHI  $< 50\%$  and/or ongoing symptoms. The secondary outcome measures were changes in daytime sleepiness, assessed by Epworth score, and changes in QOL, assessed by the eight domains in SF-36 and summary scores for physical component summary (PCS) and mental component summary (MCS).

### Statistics

Analyses were conducted before breaking the allocation code (PW). These were all intention-to-treat analyses, but per

protocol analyses were also performed according to the actual intervention received. Proportions were compared with Fischer's exact test. The outcome measures within the three intervention groups were tested by a paired non-parametric test, the Wilcoxon rank sum test.

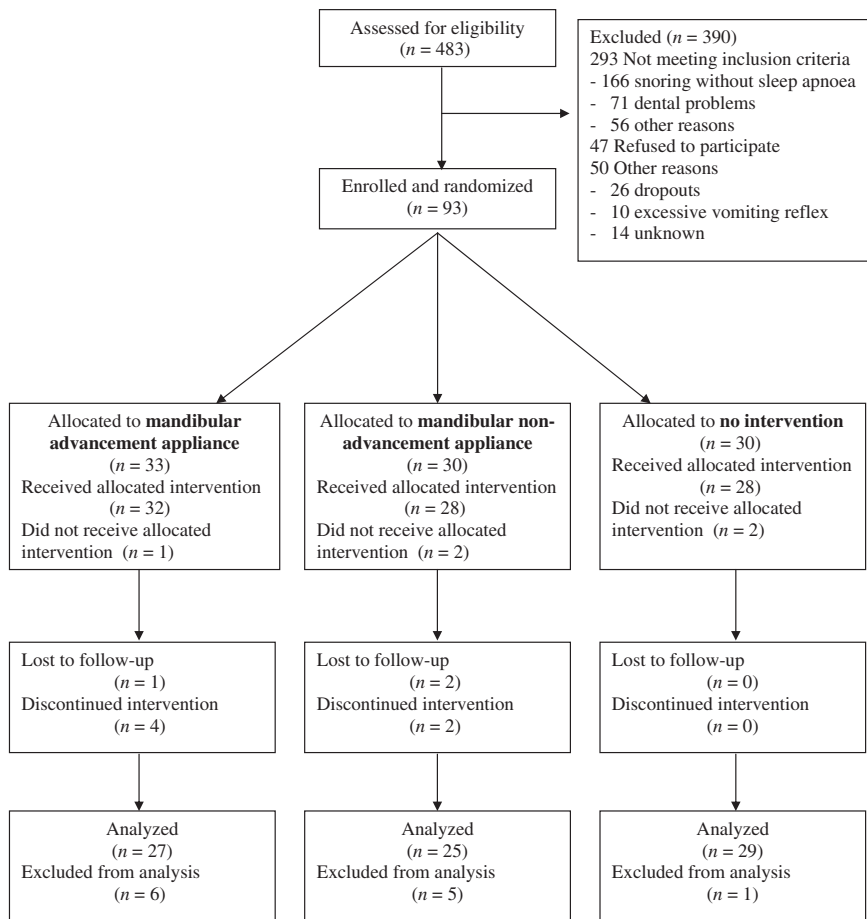
Differences in baseline values and possible confounders were controlled by a series of analyses of variance and covariance. An analysis of covariance with the baseline values as the covariate was performed to test, whether the mean values of the outcome measures differed significantly between the groups. The analysis was repeated with the stratification factors included. If the conditions for analysis of covariance were not fulfilled (as evidenced from the Shapiro-Wilk normality test and Levene's test of equality of error variances), the differences between baseline and outcome values were calculated, and a variance analysis of the differences was performed, provided the conditions were now fulfilled. If not, the mean differences were compared using the Kruskal-Wallis non-parametric analysis of variance, and *post hoc* analyses with pairwise comparisons were performed using the Bonferroni correction.

The effect of missing values was tested in a number of sensitivity analyses. In these, a missing value was replaced either by a failure or a success pattern. A failure pattern was defined by the outcome measure being equal to the baseline value, and a success pattern by the outcome measure being equal to the largest difference between the baseline and outcome values, the value being equal to zero in case of a negative difference. As the resulting distributions were somewhat artificial, the mean differences were compared using the Kruskal-Wallis test and the results compared with the corresponding analysis of covariance. The comparison between the groups was repeated using a 'worst-case' scenario, where the missing values in the MAA group were replaced by a failure pattern and in the two other groups by a success pattern. All numeric measures were expressed as mean values with standard deviations (SD) or 95% confidence intervals (CI) and statistical significance was accepted at  $P < 0.05$ .

## RESULTS

A total of 93 patients were enrolled, and were randomised as shown in Fig. 2. Of these, 81 patients (87%) completed the trial. Baseline characteristics for the patients are presented in Table 1. The enrolled patients had a mean AHI of 34.7 (95% CI 29.7–39.6), a mean BMI of 31.3 (95% CI 30.1–32.6) and a mean Epworth score of 11.0 (95% CI 10.0–11.9). The three intervention groups were well matched. There were no major differences between patients completing and patients not completing the trial.

Analyses of the primary and secondary outcome measures within the groups are presented in Table 2 and the corresponding covariance analyses in Table 3. The differences in the five SF-36 domains not listed were all non-significant. The significant differences in MCS score and the two



**Figure 2.** Profile of the trial. Among the five patients not receiving allocated intervention, two of them received the wrong intervention by mistake (one in the MNA and one in the no-intervention group got an MAA), but are included in the analyses. The remaining three were excluded: one in the MAA group dropped out, and one in the no-intervention group died before the start of intervention and one in the MNA group turned out to have periodic limb movement disorder instead of sleep apnoea.

**Table 1** Baseline characteristics of patients completing ( $n = 81$ ) and not completing ( $n = 12$ ) the trial

	<i>Mandibular advancement appliance</i> ( $n = 27$ )	<i>Mandibular non-advancement appliance</i> ( $n = 25$ )	<i>No intervention</i> ( $n = 29$ )	<i>Not completing</i> ( $n = 12$ )
Apnoea-hypopnoea index*	39.1 ± 23.8	32.6 ± 22.0	34.3 ± 26.3	29.9 ± 22.8
Sex (M/F)	23/4	20/5	23/6	10/2
Age (year)	50 ± 11	50 ± 10	49 ± 10	46 ± 6
Body mass index <sup>†</sup>	30.7 ± 5.2	31.3 ± 5.2	31.3 ± 7.4	32.9 ± 5.4
Neck circumference (cm)	42.5 ± 3.1	41.7 ± 3.6	41.8 ± 4.6	42.8 ± 3.1
Epworth sleepiness score <sup>‡</sup>	11.7 ± 4.3	10.8 ± 4.6	10.7 ± 4.6	10.1 ± 3.6
Physical component score (SF-36) <sup>§</sup>	45.5 ± 9.5	48.1 ± 9.2	46.6 ± 9.6	48.1 ± 10.8
Mental component score (SF-36) <sup>§</sup>	47.2 ± 8.5	48.8 ± 10.0	50.2 ± 8.9	52.5 ± 7.4

The values are given as mean ± SD.

\*The number of apnoeas and hypopnoeas per hour of sleep.

<sup>†</sup>The body mass index is the weight in kg divided by the square of the height in m.

<sup>‡</sup>The Epworth Sleepiness Scale ranges from 0 to 24, with 24 as the highest score of sleepiness.

<sup>§</sup> SF-36 is a Short-Form questionnaire (18 questions about physical health and 18 about mental); the higher the score, the better the health.

domains, general and mental health, in the MAA group was not found again in the covariance analysis, whereas the AHI, Epworth score and vitality domain still differed significantly between the intervention groups. Pair-wise comparisons confirmed, that the means of AHI, Epworth score and vitality in the MAA group differed significantly from that in the MNA group and no-intervention group, whereas the mean values in the latter two groups did not

differ significantly. Similar results were obtained with per protocol analyses, and when the stratification factors were included in the analysis.

Outcome measures were lacking in 12 patients, who did not complete the trial. They were included in the sensitivity analyses, and the result is shown in Table 4. Even in the 'worst-case' scenario, where patients allocated to MAA had their missing values replaced by a failure pattern, and patients in the

**Table 2** Outcome measures within the intervention groups for patients completing the trial ( $n = 81$ )

	Mandibular advancement appliance ( $n = 27$ )			Mandibular non-advancement appliance ( $n = 25$ )			No intervention ( $n = 29$ )		
	Baseline	Outcome	<i>P</i>	Baseline	Outcome	<i>P</i>	Baseline	Outcome	<i>P</i>
AHI*	39.1 ± 23.8	25.0 ± 27.5	<0.001	32.6 ± 22.0	31.7 ± 25.0	0.69	34.3 ± 26.3	33.3 ± 25.3	0.67
ESS†	11.7 ± 4.3	8.4 ± 4.3	<0.001	10.8 ± 4.6	9.6 ± 4.2	0.050	10.7 ± 4.6	10.0 ± 4.8	0.39
PCS‡ (SF-36)	45.5 ± 9.5	46.5 ± 8.0	0.21	48.1 ± 9.2	47.5 ± 11.2	0.40	46.6 ± 9.6	47.3 ± 8.7	0.69
MCS‡ (SF-36)	47.2 ± 8.5	51.1 ± 8.0	0.039	48.8 ± 10.0	49.8 ± 8.5	0.48	50.2 ± 8.9	51.2 ± 7.8	0.79
General health‡	60.7 ± 21.9	66.7 ± 20.8	0.011	66.6 ± 22.4	66.0 ± 22.1	0.94	62.7 ± 19.8	67.0 ± 19.5	0.08
Mental health‡	71.0 ± 14.7	76.4 ± 13.8	0.016	78.4 ± 19.5	80.4 ± 12.9	0.54	79.6 ± 15.2	79.0 ± 15.4	0.61
Vitality‡	41.5 ± 23.4	59.4 ± 24.7	<0.001	47.8 ± 26.7	47.0 ± 26.4	1.0	48.1 ± 24.3	51.3 ± 23.4	0.44

The values are given as the mean ± SD. The sleep variables were tested by a paired non-parametric test, the Wilcoxon test.

\*Apnoea–hypopnoea index, the number of apnoeas and hypopnoeas per hour of sleep.

†Epworth sleepiness score, the scale ranges from 0 to 24, with 24 as the highest score of sleepiness.

‡Physical Component Summary (PCS) and Mental Component Summary (MCS) and the three domains, General Health, Mental Health and Vitality in SF-36, a Short-Form questionnaire; the higher the score, the better the health.

**Table 3** Results of covariance analysis of the primary and secondary outcome measures for patients completing the trial ( $n = 81$ )

	Mandibular advancement appliance ( $n = 27$ )	Mandibular non-advancement appliance ( $n = 25$ )	No intervention ( $n = 29$ )	<i>P</i> of difference between groups
Primary outcome				
Apnoea–hypopnoea index	25.0 ± 27.5	31.7 ± 25.0	33.3 ± 25.3	<0.001
Secondary outcomes				
Epworth sleepiness score*	8.4 ± 4.3	9.6 ± 4.3	10.0 ± 4.8	0.044
Physical Component Summary (SF-36)†	46.5 ± 8.0	47.5 ± 11.2	47.3 ± 8.7	0.38
Mental Component Summary (SF-36)†	51.1 ± 8.0	49.8 ± 8.5	51.2 ± 7.8	0.20
General Health†	66.7 ± 20.8	66.0 ± 22.1	67.0 ± 19.5	0.13
Mental Health†	76.4 ± 13.8	80.4 ± 12.9	79.0 ± 15.4	0.11
Vitality†	59.4 ± 24.7	47.0 ± 26.4	51.3 ± 23.4	0.001

The values are given as the mean ± SD.

\*The Epworth Sleepiness Scale ranges from 0 to 24, with 24 as the highest score of sleepiness.

†SF-36, a Short-Form questionnaire, 18 questions about physical health and 18 about mental health, grouped in eight domains, four in physical component and one of them the domain General Health, and four in mental component, two of them the domains Mental Health and Vitality; the higher the score, the better the health.

**Table 4** Sensitivity analysis comparing results of analyses with missing values ( $n = 93$ )

Scenario	<i>P</i> of difference between intervention groups	
	Apnoea–hypopnoea index	Epworth sleepiness score
Covariance analysis	<0.001	0.044
Nonparametric analysis of variance (mean differences)	0.002	0.005
Worst case* (mean differences)	0.035	0.042

\*Missing values have been replaced by failure patterns in MAA group and success patterns in the two other groups.

two other groups had them replaced by a success pattern, the result was significant. Corresponding per protocol analyses gave similar results.

Polysomnographic results are summarised in Table 5. The MAA also had a significant effect on AHI, calculated separately for the supine and non-supine sleeping positions. Especially important, the effect was not caused by reduced time sleeping in a supine position. Further, there has been a clear shift from apnoeas to hypopnoeas. The slightest effect was seen on the sleep stages with a tendency to less stage 1 sleep and more stages 3–4.

Table 6 shows that MAA resulted in complete response for 8/27 patients (30%), all having a baseline AHI > 10 and a reduction in AHI ≥ 75%, and half of them having severe OSA with a baseline AHI > 30. Partial response was achieved for 4/27 patients (15%), whereas failure occurred for 15/27 patients (56%), although four of the patients reported resolution or improvement of clinical symptoms. With a baseline AHI > 53, none of the patients obtained larger than 50% reduction in AHI.

**Table 5** Polysomnographic results at all patients completing the trial ( $n = 81$ )

	Mandibular advancement appliance ( $n = 27$ )			Mandibular non-advancement appliance ( $n = 25$ )			No intervention ( $n = 29$ )		
	Baseline	Outcome	<i>P</i>	Baseline	Outcome	<i>P</i>	Baseline	Outcome	<i>P</i>
Apnoea-hypopnoea index (AHI)	39.1 ± 23.8	25.0 ± 27.5	<0.001	32.6 ± 22.0	31.7 ± 25.0	0.69	34.3 ± 26.3	33.3 ± 25.3	0.67
Supine AHI	56.2 ± 29.7	35.8 ± 31.9	0.008	55.2 ± 28.0	45.0 ± 30.0	0.043	44.7 ± 30.8	55.0 ± 31.3	0.26
Non-supine AHI	31.5 ± 27.2	20.3 ± 27.1	0.004	18.5 ± 23.1	25.6 ± 24.5	0.021	27.8 ± 28.6	25.0 ± 26.5	0.70
Supine time %/total sleep time	33.7 ± 18.1	37.9 ± 19.9	0.37	31.9 ± 22.6	33.8 ± 25.6	0.94	26.9 ± 22.6	32.4 ± 20.8	0.25
Apnoeas %	34.7 ± 22.9	24.5 ± 22.7	0.008	34.8 ± 25.8	37.2 ± 29.8	0.94	28.5 ± 22.4	31.8 ± 23.4	0.12
Hypopnoeas %	65.3 ± 22.9	75.5 ± 22.7	0.008	65.2 ± 25.8	62.8 ± 29.8	0.94	71.5 ± 22.4	68.2 ± 23.4	0.12
Apnoeas (s)	24.5 ± 9.6	19.7 ± 10.0	0.005	21.7 ± 8.1	21.0 ± 12.7	0.52	21.2 ± 7.4	20.6 ± 8.0	0.57
Hypopnoeas (s)	26.4 ± 7.5	24.2 ± 7.1	0.10	25.8 ± 6.7	25.6 ± 7.2	0.94	24.9 ± 6.4	24.1 ± 5.7	0.63
Mean oxygen desaturation (%)	7.4 ± 2.7	6.3 ± 2.2	0.004	6.9 ± 2.4	7.2 ± 2.7	0.86	7.0 ± 2.5	6.9 ± 2.7	0.68
Sleep stage 1%	31.9 ± 21.3	24.3 ± 16.9	0.014	32.4 ± 23.3	30.5 ± 22.4	0.80	31.2 ± 25.6	32.2 ± 22.6	0.91
Sleep stage 2%	32.5 ± 13.2	35.4 ± 10.2	0.29	31.3 ± 14.7	29.9 ± 15.8	0.78	31.4 ± 16.4	32.3 ± 14.7	0.71
Sleep stage 3%	7.0 ± 5.0	9.1 ± 5.1	0.016	9.9 ± 5.1	9.1 ± 5.6	0.57	10.3 ± 6.9	7.9 ± 5.5	0.046
Sleep stage 4%	10.2 ± 10.5	12.4 ± 9.5	0.28	9.6 ± 10.3	13.0 ± 9.6	0.043	9.5 ± 8.3	9.0 ± 10.3	0.81
REM sleep %	18.4 ± 5.6	18.7 ± 4.7	0.79	16.8 ± 7.5	17.5 ± 6.5	0.88	17.5 ± 6.2	18.5 ± 6.8	0.36

The values are given as the mean ± SD. The variables were tested by the Wilcoxon signed rank test.

**Table 6** Different success criteria with the mandibular advancement appliance (MAA)

	All MAA patients, range 10.6–111.7 ( $n = 27$ )	Baseline AHI 10 < AHI < 30, range 10.6–29 ( $n = 12$ )	AHI > 30, range 30.7–111.7 ( $n = 15$ )
Outcome in apnoea-hypopnoea index (AHI) + symptoms (snoring and sleepiness)			
Reduction in AHI ≥ 50%	13/27 patients (48)	6/12 patients (50)	7/15 patients (46)
AHI < 10	11/27 patients (40)	7/12 patients (58)	4/15 patients (26)
AHI < 5 + symptom resolution	8/27 patients (29)	4/12 patients (33)	4/15 patients (26)
Reduction in AHI ≥ 50% + AHI ≥ 5	5/27 patients (18)	2/12 patients (16)	3/15 patients (20)
Reduction in AHI ≥ 50% + AHI ≥ 5 + symptom improvement	4/27 patients (15)	1/12 patients (8)	3/15 patients (20)

Values in parentheses are percentage.

Four patients in the MAA group, two in the MNA group and none in the no-intervention group discontinued intervention because of adverse effects ( $P = 0.74$  and  $0.10$  respectively). In the MAA group, two patients could not tolerate the appliance, one patient suffered loosening of the teeth and one suffered pain of the temporomandibular joint. In the MNA group, one patient could not tolerate the appliance and one suffered loosening of the teeth.

## DISCUSSION

The main results of our trial were that MAA not only significantly reduced respiratory disturbances, but also that the effects were clinically relevant: there was a significant reduction in daytime sleepiness and increase in QOL regarding vitality. While recognizing the problem of drawing conclusions on a material broken down into small subgroups, our trial shows that MAA works across the spectrum of OSA severity until an upper limit of severity is reached – in this trial an AHI of 53.

Furthermore, no significant effect or difference between MNA and no intervention was found. Thus, it is concluded that MNA neither increases pharyngeal muscle tension, nor has a placebo effect. Consequently, our findings support the hypothesis that the mechanism of action is through protrusion of the mandible only.

Looking at the reduction in mean AHI, our result with MAA, a mean AHI reduction of 14.1 (95% CI 7.4–20.8) is comparable with most other randomised trials using sham OA as control (Gotsopoulos *et al.*, 2002; Hans *et al.*, 1997; Johnston *et al.*, 2002; Mehta *et al.*, 2001). One trial similar to ours with a parallel-group design and severe OSA patients included (Blanco *et al.*, 2005) presented outstanding results with a reduction in mean AHI twice as high as in our trial and markedly higher than the CPAP effect in well-performed, randomised, controlled trials. However, the sample size was small, the attrition rate relatively high and the effect even of the sham OA highly significant too.

In randomised trials using CPAP as control (Barnes *et al.*, 2004; Engleman *et al.*, 2002; Ferguson *et al.*, 1996; Lam *et al.*,

2007; Tan *et al.*, 2002), the reductions in mean AHI were invariably higher with CPAP than with OA, varying from 14 to 21.

Looking at the percentage of patients reaching a specific AHI target, for instance AHI lower than 5 or 10, the results are a little better in two sham OA controlled trials using a crossover design (Gotsopoulos *et al.*, 2002; Mehta *et al.*, 2001) and having complete response at 38% and 36% respectively. However, these trials have pooled the first and second period data making them non-comparable with parallel-group data. In a Cochrane systematic review of randomised placebo-controlled trials of OAs (Lim *et al.*, 2006), first period data from the crossover trials were made available to the authors of the review upon request. When these data were pooled with the parallel-group data, a significant weighted mean difference (WMD) in AHI reduction of 10.8 (95% CI 6.0–15.5) was found, whereas the combined scores from the crossover trials gave a better result, a WMD of 15.2 (95% CI 10.9–19.4). In a Nordic meta-analysis of randomised placebo-controlled trials of OAs, all crossover (The Nordic National Health Services 2007), the authors found a significant WMD in AHI reduction of 9.8 (95% CI 4.2–15.5). For comparison, we found a significant mean difference in AHI between MAA and the two control groups pooled together of 13.1 (95% CI 5.7–20.6).

The beneficial effect of MAA on excessive daytime sleepiness found in our trial, with a mean reduction in Epworth score of 3.3 (95% CI 1.8–4.8), is comparable with most other randomised trials, independent of design or type of control group. Yet again, one trial (Blanco *et al.*, 2005) has presented much better results with a mean reduction in Epworth score even higher than normally obtained with CPAP. In spite of these remarkable results on AHI and Epworth score, no significant effect on QOL, measured by SF-36, could be found. As in the case of AHI, the effect of CPAP was better in all trials comparing CPAP and OA (Engleman *et al.*, 2002; Lam *et al.*, 2007; Tan *et al.*, 2002) except in one trial (Barnes *et al.*, 2004), where the mean reduction in Epworth score was below half the value in our trial and not better in the CPAP group. In two meta-analyses of OAs (Lim *et al.*, 2006; The Nordic National Health Services 2007), the authors found a significant WMD in Epworth score reduction of 2.1 (95% CI 3.8–0.1) and 1.1 (95% CI 2.1–0.1) respectively. For comparison, we found a significant mean difference in Epworth score between MAA and the two control groups pooled together of 2.5 (95% CI 1.1–4.0).

SF-36 is among the most used and best validated questionnaires about QOL, but it is not especially appropriate for measuring OSA treatment effects, as it contains no direct questions about sleep. However, the subscale domain 'vitality' contains the most relevant questions. Among the four other randomised trials measuring QOL (Engleman *et al.*, 2002; Barnes *et al.*, 2004; Blanco *et al.*, 2005; Lam *et al.*, 2007), two of them found no significant effect of OAs on SF-36, whereas one trial (Barnes *et al.*, 2004) found significant effect of both CPAP and OAs in mean SF-36 score, and the other (Lam *et al.*, 2007) found significant effect of CPAP in six of the eight SF-36

domains and of OAs in three of them, with vitality represented in both cases.

In a Cochrane systematic review of CPAP for OSA (Giles *et al.*, 2006), the authors found a significant effect in the two SF-36 domains, physical function and general health, whereas significance in four other domains inclusive vitality was eliminated in a test for confounders. In the Nordic meta-analysis of CPAP and OA (The Nordic National Health Services 2007), the authors found a non-significant WMD in the SF-36 vitality domain of 8.0 (95% CI –1.0, 17.0) and 0.6 (95% CI –5.4, 6.6) respectively. For comparison, we found a significant mean difference in vitality between MAA and the two control groups pooled together of 17.1 (95% CI 8.4–25.9).

Schmidt-Nowara *et al.* (1995) found that the different designs of OAs had a remarkably consistent effect. In our trial there was no correlation at all between the final percentage protrusion of the mandible and reduction in AHI. This is in agreement with the findings of Gotsopoulos *et al.* (2002). Some trials have reported a placebo effect also of an inactive OA (Gotsopoulos *et al.*, 2002; Johnston *et al.*, 2002), but only on subjective outcomes.

Normally, two-piece adjustable MAAs are used in our clinic, but in our trial we needed to be certain that the patients could themselves make no adjustments to the protrusion. Further, the MAA and the MNA one-piece appliances used by us were identical in size and construction, thus resulting in a more reliable blinding. The disadvantage was a longer acclimatisation period, because the appliance had to be rebuilt for each adjustment.

Only one other placebo-controlled trial has used a sham OA with the same appearance as the active appliance (Blanco *et al.* 2005), all the other trials have used an appliance either in the upper or lower jaw as control.

Our trial has a number of strengths. We used central randomisation, which is a safeguard against allocation bias (Kjaergard *et al.*, 2001; Moher *et al.*, 1998; Schulz *et al.*, 1995). The stratified minimisation secured prognostic balance between the three intervention groups. The parallel-group design is easy to understand: both benefits and adverse effects can readily be connected with the type of intervention given. This is not so with the crossover design. Crossover trials are popular, because they reduce the number of participants needed. However, the crossover design requires both a stable disease and a reversible intervention due to some inherent deficiencies in this design: failure of the participants to return to their baseline state before the crossover, non-uniform carry-over effects, time-dependent outcome measures and negative correlation between intervention responses (Altman *et al.*, 2001; Cleophas and Zwinderman, 2002; Gluud, 2006). However, most crossover trials found rather similar effects when compared with parallel-group trials, but meta-analyses indicate, that crossover trials may overestimate the effect on respiratory disturbances. Finally, we used a control group with no intervention, which made the investigation for placebo effect and mechanism of action more powerful. To our knowledge,

no other randomised trial on this topic has used a no-intervention group as control.

Two patients were erroneously given the wrong intervention, but we conducted both intention-to-treat and per protocol analyses, and we took into account possible consequences of missing values by performing sensitivity analyses, in all cases reaching the same conclusions. However, the results obtained, using a non-parametric analysis of the delta values of Epworth score, were more significant ( $P = 0.005$ ) than those of the corresponding covariance analysis ( $P = 0.044$ ). Thus, performing an analysis of the delta values instead of performing the proper covariance analysis may have introduced some bias in favour of finding a difference between the intervention groups (Altman and Bland, 1996). The same bias may therefore be present in the non-parametric analysis of the 'worst-case' scenario, why the result of the sensitivity analysis of the Epworth score should be tempered accordingly. We did not fully reach the projected sample size, but this would only be a problem, if no significant differences between the intervention groups were observed. Sensitivity analyses, in which the projected sample size was used, confirmed the significant effects, even in the 'worst-case' scenario.

The current consensus is that CPAP offers the most effective treatment of OSA, but the present and other trials show that MAAs can be a good alternative in some cases. In a Cochrane systematic review of parallel and first period data from crossover trials of CPAP versus control, the authors found a WMD in AHI reduction of 17.0 (95% CI 14.8–19.3) and a WMD in Epworth score reduction of 3.8 (95% CI 2.5–5.2) (Giles *et al.*, 2006). In the Nordic meta-analysis of CPAP, the authors in randomised placebo-controlled trials found a WMD in AHI reduction of 13.0 (95% CI 8.3–17.7) and a WMD in Epworth score reduction of 3.0 (95% CI 1.9–4.1) (The Nordic National Health Services 2007). The trials in the first review were dominated by severe OSA and in the second by mild to moderate OSA. These results are immediately not markedly better than those obtained with OAs. However, in our trial the patients were relatively less symptomatic compared with many randomised, controlled trials of OSA treatment. For instance, the baseline mean AHI, BMI and Epworth score were all lower compared with the same baseline values in the two Cochrane systematic reviews, having a mean AHI about 41–46, BMI about 33–34 and Epworth score about 12–14. Further, WMD is not saying much about treatment efficacy on subgroups, and these meta-analyses most probably underestimate the efficacy of CPAP on very severe OSA patients.

Although it is obvious that CPAP is the treatment of choice, MAAs have their place in the treatment of OSA, but it is still a problem to select patients likely to benefit from a MAA. In a comprehensive systematic review of OAs (Ferguson *et al.*, 2006), the authors concluded, that OAs are not indicated as first-line therapy in case of severe OSA, severe daytime sleepiness or very low oxygen saturation during sleep. The American Academy of Sleep Medicine (2006) recommends OA therapy only in patients with mild to moderate OSA, who

prefer OAs to CPAP or have problems with CPAP. However, in the present trial as well as in other trials (Johnston *et al.*, 2002; Mehta *et al.*, 2001; Gotsopoulos *et al.*, 2002), MAAs have been effective even in some patients with severe OSA. One randomised parallel-group trial, comparing two MAAs with 50% and 75% fixed protrusion of the mandible, has presented extremely fine results with both appliances in very severe OSA (Walker-Engström *et al.*, 2003). The results of that study are even better than results reported from another remarkable OA trial (Blanco *et al.*, 2005) and with both appliances even better than CPAP. However, the trial was not placebo controlled. Nevertheless, from our experience only CPAP works in case of very severe OSA, and contrary to Mehta *et al.* (2001), it is not our opinion that a more aggressive advancement of the mandible may improve the outcome of MAAs, as we have chosen the maximum comfortable limit of advancement, and the adverse effects of the appliances were not a minor problem.

Even though increasing evidence of MAAs being effective for OSA, additional trials are still needed, especially for the identification of factors that may predict the outcome.

## ACKNOWLEDGMENTS

The trial was supported financially by Østdansk Sundhedsvidenskabelig Forskningsforum, Lægevidenskabelige Forskningsfond Region 3, Edith og Henrik Henriksens Mindelegat, Direktør Ib Henriksens Fond, Dansk Tandlægeforenings Fond, Direktør E. Danielsen & Hustrus Fond and Copenhagen Trial Unit, Centre for Clinical Intervention Research. The sponsors had no role in the conception or design of the trial, collection of data, analyses, interpretation or drafting of the manuscript. We thank dental technician Anne-Birthe Myrholt for skilful assistance in making the appliances.

## REFERENCES

- Altman, D. G. and Bland, J. M. Statistics notes: comparing several groups using analysis of variance. *Br. Med. J.*, 1996, 312: 1472–1473.
- Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., Gøtzsche, P. C. and Lang, T. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann. Intern. Med.*, 2001, 134: 663–694.
- American Academy of Sleep Medicine. Practice parameters for the treatment of snoring and obstructive sleep apnea with oral appliances: an update for 2005. *Sleep*, 2006, 29: 240–243.
- Barnes, M., McEvoy, R. D., Banks, S., Tarquinio, N., Murray, C. G., Vowles, N. and Pierce, R. J. Efficacy of positive airway pressure and oral appliance in mild to moderate obstructive sleep apnea. *Am. J. Respir. Crit. Care Med.*, 2004, 170: 656–664.
- Blanco, J., Zamarrón, C., Abeleira Pazos, M. T., Lamela, C. and Suarez Quintanilla, D. Prospective evaluation of an oral appliance in the treatment of obstructive sleep apnea syndrome. *Sleep Breath*, 2005, 9: 20–25.
- Cleophas, T. J. M. and Zwinderman, A. H. Crossover studies with continuous variables: power analyses. *Am. J. Ther.*, 2002, 9: 69–73.
- Engleman, H. M., McDonald, J. P., Graham, D., Lello, G. E., Kingshott, R. N., Coleman, E. L., Mackay, T. W. and Douglas, N. J. Randomized crossover trial of two treatments for sleep



- apnea/hypopnea syndrome. *Am. J. Respir. Crit. Care Med.*, 2002, 166: 855–859.
- Ferguson, K. A., Takashi, O., Lowe, A. A., Keenan, S. P. and Fleetham, J. A. A randomized crossover study of an oral appliance vs nasal-continuous positive airway pressure in the treatment of mild-moderate obstructive sleep apnea. *Chest*, 1996, 109: 1269–1275.
- Ferguson, K. A., Cartwright, R., Rogers, R. and Schmidt-Nowara, W. Oral appliances for snoring and obstructive sleep apnea: a review. *Sleep*, 2006, 29: 244–262.
- Fujita, S., Conway, W., Zorick, F. and Roth, T. Surgical correction of anatomic abnormalities in obstructive sleep apnea syndrome: uvulopalatopharyngoplasty. *Otolaryngol. Head Neck Surg.*, 1981, 89: 923–934.
- George, P. T. A new instrument for functional appliance bite construction. *J. Clin. Orthod.*, 1992, 26: 721–723.
- Giles, T. L., Lasserson, T. J., Smith, B. J., White, J., Wright, J. and Cates, C. J.. Continuous positive airways pressure for obstructive sleep apnoea in adults. *Cochrane Database Syst. Rev.*, 2006, Art.No.: CD001106. DOI: 10.1002/14651858.CD001106.pub3.
- Gluud, C. The culture of designing hepato-biliary randomised trials. *J. Hepatol.*, 2006, 44: 607–615.
- Gotsopoulos, H., Chen, C., Qian, J. and Cistulli, P. A. Oral appliance therapy improves symptoms in obstructive sleep apnea. *Am. J. Respir. Crit. Care Med.*, 2002, 166: 743–748.
- Hans, M. G., Nelson, S., Luks, V. G., Lorkovich, P. and Baek, S. J. Comparison of two dental devices for treatment of obstructive sleep apnea syndrome (OSAS). *Am. J. Orthod. Dentofacial. Orthop.*, 1997, 111: 562–570.
- Johns, M. W. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep*, 1991, 14: 540–545.
- Johnston, C. D., Gleadhill, I. C., Cinnamon, M. J., Gabbey, J. and Burden, D. J. Mandibular advancement appliances and obstructive sleep apnoea: a randomized clinical trial. *Eur. J. Orthod.*, 2002, 24: 251–262.
- Kapur, V., Blough, D. K., Sandblom, R. E., Hert, R., de Maine, J. B., Sullivan, S. D. and Psaty, B. M. The medical cost of undiagnosed sleep apnea. *Sleep*, 1999, 22: 749–755.
- Kjaergard, L. L., Villumsen, J. and Gluud, C. Reported methodological quality and discrepancies between large and small randomised trials in meta-analyses. *Ann. Intern. Med.*, 2001, 135: 982–989.
- Lam, B., Sam, K., Mok, W. Y. W., Cheung, M. T., Fong, D. Y. T., Lam, J. C. M., Lam, D. C. L., Yam, L. Y. C. and Ip, M. S. M. Randomised study of three non-surgical treatments in mild to moderate obstructive sleep apnoea. *Thorax*, 2007, 62: 354–359.
- Lavie, P., Herer, P. and Hoffstein, V. Obstructive sleep apnoea syndrome as a risk factor for hypertension: population study. *Br. Med. J.*, 2000, 320: 479–482.
- Lim, J., Lasserson, T. J., Fleetham, J. and Wright, J.. Oral appliances for obstructive sleep apnoea. *Cochrane Database Syst. Rev.*, 2006, Art. No.: CD004435. DOI: 10.1002/14651858.CD004435.pub3.
- Marin, J. M., Carrizo, S. J., Vicente, E. and Agusti, A. G. N. 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: 1046–1053.
- McArdle, N., Devereux, G., Heidarnjad, H., Engleman, H. M., Mackay, T. W. and Douglas, N. J. Long-term use of CPAP therapy for sleep apnea/hypopnea syndrome. *Am. J. Respir. Crit. Care Med.*, 1999, 159: 1108–1114.
- Mehta, A., Qian, J., Petocz, P., Darendeliler, M. A. and Cistulli, P. A. A randomized, controlled study of a mandibular advancement splint for obstructive sleep apnea. *Am. J. Respir. Crit. Care Med.*, 2001, 163: 1457–1461.
- Moher, D., Pham, B., Jones, A., Cook, D. J., Jadad, A. R., Moher, M., Tugwell, P. and Klassen, T. P. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet*, 1998, 352: 609–613.
- Peker, Y., Hedner, J., Johansson, A. and Bende, M. Reduced hospitalization with cardiovascular and pulmonary disease in obstructive sleep apnea patients on nasal CPAP treatment. *Sleep*, 1997, 20: 645–653.
- 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: 159–165.
- Peppard, P. E., Young, T., Palta, M. and Skatrud, J. Prospective study of the association between sleep-disordered breathing and hypertension. *N. Engl. J. Med.*, 2000, 342: 1378–1384.
- Popescu, G., Latham, M., Allgar, V. and Elliott, M. W. Continuous positive airway pressure for sleep apnoea/hypopnoea syndrome: usefulness of a 2 week trial to identify factors associated with long term use. *Thorax*, 2001, 56: 727–733.
- Rechtschaffen, A. and Kales, A. (Eds) *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. Brain Information Service/Research Institute, UCLA, Los Angeles, CA, 1968.
- Schmidt-Nowara, W., Lowe, A., Wiegand, L., Cartwright, R., Perez-Guerra, F. and Menn, S. Oral appliances for the treatment of snoring and obstructive sleep apnea: a review. *Sleep*, 1995, 18: 501–510.
- Schulz, K. F., Chalmers, I., Hayes, R. J. and Altman, D. G. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*, 1995, 273: 408–412.
- Sullivan, C. E., Issa, F. G., Berthon-Jones, M. and Eves, L. Reversal of obstructive sleep apnoea by continuous positive airway pressure applied through the nares. *Lancet*, 1981, 1: 862–865.
- Sundaram, S., Bridgman, S. A., Lim, J. and Lasserson, T. J.. Surgery for obstructive sleep apnoea. *Cochrane Database Syst. Rev.*, 2005, Art. No.: CD001004. DOI: 10.1002/14651858.CD001004.pub2.
- Talmant, J., Renaudin, S. and Renaud, P. Ventilation and mechanics of the oropharynx. *Rev. Orthop. Dento Faciale*, 1998, 32: 105–166.
- Tan, Y. K., L'Estrange, P. R., Luo, Y. M., Smith, C., Grant, H. R., Simonds, A. K., Spiro, S. G. and Battagel, J. M. Mandibular advancement splints and continuous positive airway pressure in patients with obstructive sleep apnoea: a randomized cross-over trial. *Eur. J. Orthod.*, 2002, 24: 239–249.
- The Nordic National Health Services. *Obstructive Sleep Apnoea Syndrome: A Systematic Literature Review*. 2007. Available at: [http://www.sbu.se/upload/Publikationer/Content0/1/som-napne\\_fulltext.pdf](http://www.sbu.se/upload/Publikationer/Content0/1/som-napne_fulltext.pdf).
- Tsuiki, S., Lowe, A. A., Almelda, F. R. and Fleetham, J. A. Effects of an anteriorly titrated mandibular position on awake airway and obstructive sleep apnea severity. *Am. J. Orthod. Dentofacial. Orthop.*, 2004, 125: 548–555.
- Waldhorn, R. E., Herrick, T. W., Nguyen, M. C., O'Donnell, A. E., Sodero, J. and Potolicchio, S. J. Long-term compliance with nasal continuous positive airway pressure therapy of obstructive sleep apnea. *Chest*, 1990, 97: 33–38.
- Walker-Engström, M. L., Ringqvist, I., Vestling, O., Wilhelmsson, B. and Tegelberg, Å. A prospective randomized study comparing two different degrees of mandibular advancement with a dental appliance in treatment of severe obstructive sleep apnea. *Sleep Breath*, 2003, 7: 119–130.
- Ware, J. E.. The SF-36 health survey. In: B. Spiker (Ed.) *Quality of Life and Pharmaco-economics in Clinical Trials*, 2nd edn. Lippincott-Raven Publishers, Philadelphia, PA, 1996: 337–345.
- Yaggi, H. K., Concato, J., Kernan, W. N., Lichtman, J. H., Brass, L. M. and Mohsenin, V.. Obstructive sleep apnea as a risk factor for stroke and death. *N. Engl. J. Med.*, 2005, 353: 2034–2041.
- Young, T., Palta, M., Dempsey, J., Skatrud, J., Weber, S. and Badr, S. The occurrence of sleep-disordered breathing among middle-aged adults. *N. Engl. J. Med.*, 1993, 328: 1230–1235.