IQWiG Reports – Commission No. N18-03

Mandibular advancement device in mild to moderate obstructive sleep apnoea in adults[1]

Extract

[1] Translation of Chapters 1 to 6 of the final report N18-03 *Unterkieferprotrusionsschiene bei leichter bis mittelgradiger obstruktiver Schlafapnoe bei Erwachsenen* (Version 2.0; Status: 7 May 2020 [German original], 10 June 2020 [English translation]). Please note: This document was translated by an external translator and is provided as a service by IQWiG to English-language readers. However, solely the German original text is absolutely authoritative and legally binding.

Publishing details

Publisher:
Institute for Quality and Efficiency in Health Care

Topic:
Mandibular advancement device in mild to moderate obstructive sleep apnoea in adults

Commissioning agency:
Federal Joint Committee

Commission awarded on:
13 September 2018

Internal Commission No.:
N18-03

Address of publisher:
Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln
Germany

Phone: +49 221 35685-0
Fax: +49 221 35685-1
E-mail: berichte@iqwig.de
Internet: www.iqwig.de

This report was prepared in collaboration with external experts.

The responsibility for the contents of the report lies solely with IQWiG.

According to §139b (3) No. 2 of Social Code Book (SGB) V, Statutory Health Insurance, external experts who are involved in the Institute's research commissions must disclose "all connections to interest groups and contract organizations, particularly in the pharmaceutical and medical devices industries, including details on the type and amount of any remuneration received". The Institute received the completed *Form for disclosure of potential conflicts of interest* from each external expert. The information provided was reviewed by a Committee of the Institute specifically established to assess conflicts of interests. The information on conflicts of interest provided by the external experts and external reviewers is presented in Chapter A8 of the full report. No conflicts of interest were detected that could endanger professional independence with regard to the work on the present commission.

External experts

- Boris A. Stuck, Clinic for Ear, Nose and Throat Medicine, University Hospital Gießen and Marburg GmbH, Marburg site, Germany

IQWiG thanks the external expert for his collaboration in the project.

IQWiG employees

- Martina Lietz
- Konstanze Angelescu (until 10/2019)
- Marco Knelangen
- Julia Kreis (from 11/2019)
- Fabian Lotz
- Markus von Pluto Prondzinski

Keywords: Mandibular Advancement, Sleep Apnoea – Obstructive, Benefit Assessment, Systematic Review

Key statement

Research question

The aims of this investigation are to

- assess the benefit of treatment with a mandibular advancement device (MAD therapy) and

- assess the benefit of MAD therapy in comparison with treatment using positive airway pressure applied through a mask (PAP therapy)

each in adult patients with obstructive sleep apnoea and in respect of patient-relevant outcomes.

Any concomitant conservative treatment measures should be used in the same manner in the intervention and comparator groups.

Conclusion

Research question 1 – MAD therapy versus no treatment or placebo treatment

For excessive daytime sleepiness, an indication of benefit of MAD therapy in comparison with no treatment or with placebo treatment not affecting the mandibular position was found. For fatigue, there was a hint of benefit of MADs in comparison with placebo devices.

There was no hint of any benefit or harm of MAD therapy for the outcomes of sleep quality, cognitive performance (vigilance), cognitive performance (executive functions), depressive symptoms, anxiety symptoms, psychological symptoms, somatic symptoms (headache), health-related quality of life, activities of daily living as well as participation in professional and social life, serious adverse events, and discontinuation due to adverse events.

No data are available for the outcomes of overall mortality or overall survival and cardiovascular morbidity.

In summary, for research question 1, an advantage of MAD therapy regarding the main symptom of OSA, excessive daytime sleepiness, was found in comparison to either no treatment or placebo treatment not affecting the mandibular position. This advantage is not called into question by the results for other patient-relevant outcomes. For fatigue, an advantage of MAD therapy over a placebo device was found as well.

Research question 2 – MAD therapy versus PAP therapy

For excessive daytime sleepiness, an indication of noninferiority of MAD therapy in comparison with PAP therapy was derived.

There was no hint of greater benefit or harm of MAD therapy for the outcomes of sleep quality, cognitive performance (vigilance), cognitive performance (executive functions), depressive symptoms, anxiety symptoms, psychological symptoms, health-related quality of life, activities

of daily living as well as participation in professional and social life, serious adverse events, and discontinuation due to adverse events.

No (usable) data were available for the outcomes of all-cause mortality or overall survival, somatic symptoms, and cardiovascular morbidity.

In summary, for research question 2, MAD therapy was shown to be noninferior to PAP therapy regarding the main symptom of OSA, excessive daytime sleepiness, and simultaneously, no relevant disadvantage of MAD therapy versus PAP therapy was found for other patient-relevant outcomes.

Table of contents

List of tables

Page

List of abbreviations

Abbreviation	Meaning
AHI	Apnoea-Hypopnoea Index
BAI	Beck Anxiety Inventory
BDI	Beck Depression Inventory
BQ	Berlin Questionnaire for sleep apnoea
CI	Confidence interval
COWAT	Controlled Oral Word Association Test
CPAP	Continuous positive airway pressure
DGSM	Deutsche Gesellschaft für Schlafforschung und Schlafmedizin (German Sleep Society)
ESS	Epworth Sleepiness Scale
FOSQ	Functional Outcomes of Sleep Questionnaire
G-BA	Gemeinsamer Bundesausschuss (Federal Joint Committee)
HADS	Hospital Anxiety and Depression Scale
ICSD	International Classification of Sleep Disorders
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Institute for Quality and Efficiency in Health Care)
ITT	Intention to treat
MAD	Mandibular advancement device
MD	Mean difference
MFIS	Modified Fatigue Impact Scale
MURT	Multiple Unprepared Reaction Time
NHP	Nottingham Health Profile
ODI	Oxygen desaturation index
OSA	Obstructive sleep apnoea
PAP	Positive airway pressure
PASAT	Paced Auditory Serial Addition Task
PG	Polygraphy
POMS	Profile of Mood States
PSQI	Pittsburgh Sleep Quality Index (Pittsburgh Schlafqualitätsindex)
PSG	Polysomnography
PVT	Psychomotor Vigilance Test
RCT	Randomized controlled trial
SAE	Serious adverse event
SCL-90-R	Symptom Checklist-90-Revised

Abbreviation	Meaning
SF-36	36-Item Short-Form Health Survey
SGB	Sozialgesetzbuch (Social Code Book)
UPPP	Uvulopalatopharyngoplasty

1 Background

Obstructive sleep apnoea (OSA) is characterized by recurrent partial or complete obstruction of the upper airway during sleep. In the presence of additional factors, muscle relaxation during sleep leads to the narrowing of the upper airway or even temporary airway closure because in predisposed patients, the airway can no longer be held open against the negative pressure arising during inspiration. Important causal factors include those of a functional nature (e.g. limited reflex control of pharyngeal muscle activity, fluid redistribution upon lying down, instability of respiratory control) and those of an anatomic nature (e.g. the mandible and tongue falling back posteriorly, particularly when resting in a supine position, upper airway constriction, e.g. due to retrognathia, tonsillar hypertrophy, septal deviation – or pharyngeal fat tissue in obese patients) [1-5].

OSA is characterized by reduced respiratory airflow (hypopnoea), cessation of breathing (apnoea), and greater respiratory effort due to obstruction of the upper airway and continued respiratory activity. This increased breathing effort normally leads to central nervous activation (arousal/wake-up response), which terminates the respiratory event. The affected person is often unaware of the arousals. OSA is also commonly associated with loud snoring [6,7].

OSA leads to sleep fragmentation and therefore often non-restorative sleep, which is, in turn, associated with exhaustion [8], excessive daytime sleepiness as a main symptom [1,2], involuntary dozing [1], declines in cognitive performance, and higher accident rates [9,10]. Untreated OSA is associated with hypertension, cardiovascular events such as myocardial infarction and cerebrovascular accidents as well as increased mortality rates [1].

Referring to the above-mentioned causal factors, OSA risk factors include obesity [11-13], certain craniofacial conditions such as posterior mandibular position, enlarged tonsils [11,14], smoking [15], and to a lesser degree menopause [16,17] and familial predisposition to OSA (family history) [18,19].

The Apnoea-Hypopnoea Index (AHI) is considered one of the most important diagnostic parameters and is used to diagnose OSA and determine its severity [1]. It indicates the average number of apnoeas, hypopnoeas, and arousals per hour of sleep. According to a guideline issued by the German Sleep Society (DGSM) and drawing on the criteria of the International Classification of Sleep Disorders (ICSD-3), OSA is present if the AHI shows more than 15 events (each lasting $\geq$ 10 seconds) per hour of sleep or recording period (AHI > 15 [hereinafter, AHI figures are reported per hour, in this case an AHI of over 15 per hour]) or in case of AHI $\geq$ 5 in conjunction with additional clinical symptoms (e.g. excessive daytime sleepiness) or relevant comorbidities. The respiratory disorder must not be explicable by another sleep disorder or medical condition or by medications or other substances [1].

OSA is considered mild up to AHI = 15, moderate at AHI > 15, and severe at AHI > 30 [1]; however, the severity cannot be defined solely by the number of respiratory events; symptoms and patient comorbidities should be considered as well [1,2].

OSA prevalence depends on sex and age; it is higher for men and increases with age [2]. Young et al. 1993 published U.S. prevalences on the basis of the prospective Wisconsin Sleep Cohort Study (WSCS), which started in 1988. In participants 30 to 60 years of age, the prevalence of AHI ≥ 15 was 4% for women and 9.1% for men [20]. The Swiss cohort study HypnoLaus from 2015 [21], which studied patients 35 through 75 years of age from 2009 through 2013 suggests higher OSA prevalence in this population, at prevalence rates of 83.8% (men) and 60.8% (women) for mild to severe OSA and 49.7% (men) and 23.4% (women) for moderate to severe OSA.

One reason for the generally observed rise in prevalence rates might be the fact that, over the years, measuring techniques have improved and definitions of the various respiratory events have changed, while thresholds have remained unchanged [2]. Prevalence also seems to be increasing due to the fact that the proportion of obese patients in the population has risen over time [11-13].

In accordance with the G-BA guideline, OSA diagnostic testing is done in stages [22]. Stage 1 uses a questionnaire-based medical history. Stage 2 involves a clinical examination, particularly for metabolic disorders, cardiovascular disorders, ventilation disorders, and neurological and psychiatric disorders. If OSA is suspected, this is followed by the 3rd stage, which involves differential diagnostics using cardiorespiratory polygraphy (PG). PG is an outpatient examination which is typically conducted in the affected person's home environment. It registers, for instance respiration, blood oxygenation, heart rate or pulse, and body position. Cardiorespiratory polysomnography (PSG) (stage 4) is conducted only if the examinations done in stages 1 through 3 cannot determine with sufficient certainty whether ventilation therapy by positive airway pressure (PAP), for instance continuous PAP (CPAP), is necessary [22]. PSG involves the examinations of PG plus electroencephalography (EEG), electromyography (EMG), and electrooculography (EOG). PSG is usually conducted in a sleep laboratory.

The choice of OSA therapy depends in part on the severity of disease and causal factors [4]. In case of mild OSA, conservative measures such as weight reduction, sleep hygiene measures (no alcohol, no smoking, etc.), or positional therapy (avoiding sleeping in a supine position) may suffice [4]. At higher severity levels, (C)PAP is the standard of care. In (C)PAP, positive pressure ventilation keeps the airway open [1].

According to the guideline, mild to moderate OSA can also be treated using a mandibular advancement device (MAD), which is worn during sleep [1]. Its use requires the patient to have adequate dental status, no contraindications, such as prior illnesses or temporomandibular joint dysfunction, and sufficient mandibular protrusion [2,23]. The intraoral device moves the mandible, tongue, and other structures forward. This is intended to keep the airway mechanically open by expanding the pharyngeal lumen [3,24]. This method is usually well tolerated and often preferred by patients over PAP therapy. In severe OSA, it is generally used only in case of ineligibility for or intolerance to PAP therapy [1,25]. Devices are either custom-made based on a dental impression [26] or made of thermoplastic material; in the latter case,

they are softened by heating them in a water bath, placed on the dental arch in this softened state, and moulded intraorally once a suitable degree of protrusion has been achieved (boil-and-bite devices) [27,28]. Furthermore, monoblock devices are distinguished from twinblock devices, many of which can be later readjusted in terms of the degree of mandibular protrusion [29,30].

Other treatment approaches involve surgical procedures such as the removal of hyperplastic tonsils [31], often in conjunction with uvulopalatopharyngoplasty (UPPP) [32], stimulation of the hypoglossal nerve by means of an implanted device [1,31], or maxillomandibular advancement (MMA) surgery [1,33]. UPPP and MMA are more commonly used in moderate to severe OSA. Furthermore, it is possible to combine multiple treatment approaches, including weight reduction and positional therapy [1]. Drug therapy of OSA is not recommended by the DGSM S3 guideline [1].

2 Research question

The aims of this investigation are to

- assess the benefit of treatment with a mandibular advancement device (MAD therapy) and
- assess the benefit of MAD therapy in comparison with treatment using positive airway pressure applied through a mask (PAP therapy)

each in adult patients with obstructive sleep apnoea and in respect of patient-relevant outcomes.

Any concomitant conservative treatment measures should be used in the same manner in the intervention and comparator groups.

3 Methods

The target population of the benefit assessment was adult patients with OSA. The treatment need is derived from the diagnosis of OSA. The experimental intervention was a mandibular advancement device. The comparator intervention was

- no treatment or treatment with a placebo
- PAP therapy

Any concomitant conservative treatment measures were to be used in the same manner in the intervention and comparator groups.

The investigation examined the following patient-relevant outcomes:

- Mortality
- Morbidity, particularly
  - sleep quality (e.g. severity of sleep disorder in the form of non-restorative sleep)
  - excessive daytime sleepiness, including endangerment of self or others due to excessive daytime sleepiness
  - cognitive performance
  - psychological symptoms, particularly affective symptoms
  - somatic symptoms, particularly cardiovascular events (myocardial infarction, cerebrovascular accident, etc.), symptoms of heart failure and respiratory insufficiency, symptoms of hypertension and headache
  - cardiac arrhythmia requiring treatment
- Health-related quality of life, including activities of daily living as well as participation in professional and social life
- Adverse events

AHI and the oxygen saturation index (OSI) were examined as supplementary information. However, it is not possible to derive a (greater) benefit on this basis alone.

Only randomized controlled trials (RCTs) were included in the benefit assessment. The intervention period was to be no less than 1 week.

A systematic search for primary literature was conducted in the databases MEDLINE, Embase, and Cochrane Central Register of Controlled Trials. In parallel, a search for relevant systematic reviews was conducted in the databases MEDLINE, Embase, Cochrane Database of Systematic Reviews, and HTA Database.

The following sources of information and search techniques were additionally used: trial registries, documents sent by the G-BA, reviews of reference lists, documents made available from hearing procedures, and requests to authors.

Relevant studies were selected by 3 persons independently from one another. The results of the selection were summarized after the full text assessment. Data were extracted into standardized tables. To assess the qualitative certainty of results, the risk of bias at study and outcome levels was assessed and rated as high or low. The results of the individual studies were described for each outcome.

To the extent that the studies were comparable in terms of their research questions and relevant characteristics, and no meaningful heterogeneity was observed, the results from individual studies were quantitatively combined in metaanalyses.

For each outcome, a conclusion was drawn on the evidence for (greater) benefit and (greater) harm, with 4 levels of certainty of conclusions: Proof (highest certainty of conclusions), indication (moderate certainty of conclusions), hint (lowest certainty of conclusions), or neither of the above 3. The latter was the case if no data were available or the available data did not permit classification into one of the 3 other categories. In that case, the conclusion “There is no hint of (greater) benefit or (greater) harm” was drawn.

Furthermore, for research question 2, MAD versus PAP therapy, a noninferiority test was performed regarding excessive daytime sleepiness as a main symptom of OSA. The literature states that many patients prefer a MAD over PAP therapy because the latter is often experienced as more bothersome, and MADs are also considered easier to use [1,34,35]. Noninferiority of MAD therapy in comparison with PAP therapy with regard to excessive daytime sleepiness can be considered an advantage so long as there is no relevant disadvantage compared to PAP therapy regarding other patient-relevant outcomes. No noninferiority testing was performed for any other outcomes.

4 Results

4.1 Results of the comprehensive information retrieval

The information retrieval found 37 randomized controlled studies to be relevant for the 2 research questions of this benefit assessment. A total of 8 ongoing and 2 planned studies were found. Furthermore, a total of 4 studies of unclear status and 2 completed studies without reported results were found.

The most recent search was conducted on 22 August 2019. The search strategies for bibliographic databases and trial registries are found in the appendix.

Among the included studies were 4 studies with more than 2 treatment arms; these studies provided relevant comparisons for both research questions of this report and are therefore listed for both research questions in this benefit assessment (Aarab 2011 [36], Barnes 2004 [37], Dal-Fabbro 2014 [26], and Lam 2007 [29]).

Table 1: Study pool of the benefit assessment

Study	**Available documents**	
	Full publication (in professional journals)	**Registry entry / results report from the study registries**
Research question 1 – MAD therapy versus no treatment or placebo treatment		
Aarab 2010[a]	Yes [38]	No
Aarab 2011	Yes [36,39-41]	Yes [42] / no
Andrén 2013	Yes [43]	No/no
Barnes 2004[a]	Yes [37]	No/no
Blanco 2005[b]	Yes [44]	No/no
Bloch 2000	Yes [45]	No/no
Dal-Fabbro 2014[a]	Yes [26]	No/no
Durán-Cantolla 2015[a]	Yes [46]	No/no
Gagnadoux 2017	Yes [47,48]	Yes [49] / no
Godoy 2017	Yes [50]	Yes [51] / no
Gotsopoulos 2002	Yes [52-54]	No/no
Hans 1997[b]	Yes [55]	No/no
Johnston 2002	Yes [56]	No/no
Lam 2007	Yes [29]	No/no
Marklund 2015	Yes [57,58]	Yes [59] / no
Marklund 2016[b]	Yes [60]	No/no
Mehta 2001	Yes [61]	No/no
Petri 2008	Yes [62]	Yes [63] / no
Sjöholm 1994	Yes [64]	No/no
Teixeira 2013	Yes [65]	No/no
TOMADO	Yes [66,67]	Yes [68] / no

(continued)

Table 1: Study pool of the benefit assessment (continued)

Study	Available documents	
	Full publication (in professional journals)	**Registry entry / results report from the study registries**
Research question 2 – MAD therapy versus PAP therapy		
Aarab 2011	Yes [36,39-41]	Yes [42] / no
Arya 2014	Yes [69]	No/no
Banhiran 2018	Yes [28]	Yes[c] [70,71] / no
Barnes 2004[a]	Yes [37]	No/no
COMET[b]	No	Yes [72] / yes [73]
Dal-Fabbro 2014[a]	Yes [26]	No/no
De Vries 2019[b]	Yes [74]	Yes [75] / no
El-Solh 2017	Yes [76]	Yes [77,78] / yes [79]
Engleman 2002	Yes [80]	No/no
Ferguson 1996[a]	Yes [81]	No/no
Ferguson 1997	Yes [82]	No/no
Gagnadoux 2009[a]	Yes [83,84]	Yes [85] / no
Glos 2016	Yes [86]	Yes [87] / no
Hoekema 2006	Yes [30,88-95]	Yes [96] / no
Lam 2007	Yes [29]	No/no
Phillips 2013	Yes [97,98]	Yes [99] / no
Randerath 2002[a]	Yes [27]	No/no
Schütz 2013[b]	Yes [100]	Yes [51] / yes [101]
Tan 2002	Yes [102]	No/no
Yamamoto 2019[a]	Yes [103]	Yes [104] / no

a: Crossover study disregarding data dependence structures.
b: Included only as a formality due to a lack of usable data.
c: According to the registry entry, results have been transmitted to ClinicalTrials.gov, but have yet to be published.
MAD: mandibular advancement device; PAP: positive airway pressure

4.2 Results on research question 1 – MAD therapy versus no treatment or placebo

4.2.1 Characteristics of the studies included in the assessment

For research question 1, a total of 21 studies were included, of which the results of 18 studies were used in the benefit assessment. The results of 3 studies were excluded from the benefit assessment due to unusable data: These studies involved either an excessive percentage of patients excluded from the analysis (Blanco 2005 [44]), an excessive difference between intervention groups in terms of the percentages of excluded patients (Hans 1997 [55]), or unclear patient flow and measuring times (Marklund 2016 [60]). The information provided below applies to the 18 studies with usable data.

Eleven studies had 2 arms, 6 studies had 3 arms, and 1 study had 4 arms. In each of the 2-arm studies, one MAD was compared with placebo. In 9 studies, the control group received a placebo device, and in 2 studies, a palatal plate. Among the 3-arm studies, 3 studies compared MAD therapy with PAP therapy and placebo (placebo device in 1 study, placebo tablet in 1 study, and palatal plate in 1 study). In 1 study, MAD therapy was compared with PAP therapy and conservative treatment measures, which patients in the MAD and PAP therapy arms received as accompanying therapy. In 1 study, MAD therapy was compared to placebo and no treatment. In another 3-arm study, 2 different MADs were compared with no treatment. (The benefit assessment primarily considered the Herbst appliance, while taking the monoblock device into account in a sensitivity analysis.) In the 4-arm study, 3 different MADs were compared with no treatment. (However, the benefit assessment included only 2 of these MADs, namely a custom-made device and a boil-and-bite device, i.e. the 2 device types which differ from one another as much as possible; the 3rd MAD was a hybrid, custom made on the basis of a boil-and-bite device, and was therefore disregarded.)

Seven studies used a parallel-group design, and 11 studies, a crossover design. For 12 studies, the available documents did not provide any information about the period during which they were conducted. The studies providing this information were conducted between 2003 and 2016. Between 6 and 150 patients were randomized, for a total of 1032 patients across all 18 studies. In crossover studies, each patient was entered into the calculation only once; in 3-arm studies of parallel design, only the individuals in the study arms of interest for research question 1 were included in the calculation.

In 16 studies, custom-made MADs were used, and in 1 study, a custom-made device was used in one study arm and a boil-and-bite device in another study arm (for 1 study, no information was available on the fabrication method). In 14 studies, twinblock appliances were used, in 2 studies, monoblock appliances, and in 2 studies, both monoblock and twinblock appliances. Among the studies with usable ESS data, 9 studies used twinblock appliances, and 2 studies, monoblock appliances; in 2 studies, both monoblock and twinblock appliances were used. The treatment durations were between 1 week and 18 months.

All studies except Godoy 2017 [50] included patients with OSA diagnosed on the basis of AHI or respiratory disturbance index findings. Godoy 2017 included only patients with upper airway resistance syndrome (UARS), which is now considered a form of OSA. Given the inclusion criteria of Godoy 2017 (e.g. RDI > 5 and/or airflow limitation > 30% of total sleep time, plus ESS ≥ 10 and/or MFIS ≥ 38), patients meet the OSA criteria in accordance with ICSD-3. In 15 studies, AHI was among the inclusion criteria. Two studies defined merely upper inclusion limits for AHI: One of these studies was Godoy 2017 [50], with a limit of AHI ≤ 5. The other study was Marklund 2015 [57,58]: Marklund 2015 included patients with mild to moderate OSA, for whom a limit of AHI < 30 applied, as well as patients with AHI < 5. However, patients with AHI < 5 made up less than 20% of the population of each study arm. Six studies defined both a lower AHI limit (all using AHI ≥ 5) and an upper limit (3 studies using AHI < 30 or ≤ 30; 1 study using AHI ≤ 40; 2 studies using AHI ≤ 45). Seven studies defined only a lower

inclusion limit for AHI (3 studies using AHI > 5 or ≥ 5; 2 studies using AHI ≥ 10; 1 study using AHI ≥ 20; 1 study using AHI ≥ 30).

Four studies reported why the included patients did not receive PAP therapy: Two studies stated that patients declined PAP therapy or preferred MAD therapy (Bloch 2000 [45], Petri 2008 [62]). One study reported that patients did not tolerate PAP therapy (Gagnadoux 2017 [47,48]). One study stated that PAP therapy was declined, not tolerated, or considered unnecessary (TOMADO [66,67]). In 4 further studies (Aarab 2011 [36,39-41], Barnes 2004 [37], Dal-Fabbro 2014 [26], Lam 2007 [29]), PAP therapy was administered in a 3rd study arm. The remaining 10 studies included for research question 1 did not provide any reasons why PAP therapy was not performed.

4.2.2 Overview of assessment-relevant outcomes

Data on patient-relevant outcomes were extracted from 16 studies; 2 studies provided usable data only on the non-patient-relevant outcomes of AHI (Teixeira 2013 [65]) and the oxygen desaturation index (ODI) (Sjöholm 1994 [64]). Table 2 presents an overview of the available data on patient-relevant outcomes as provided by the included studies. Four studies reported sleep quality data which were not usable for the benefit assessment. Unusable data on the outcomes of fatigue and health-related quality of life were each reported by one further study. However, usable data on the outcomes of sleep quality, fatigue, and health-related quality of life were available from other studies. None of the studies reported data on the outcomes of mortality or cardiovascular morbidity.

Data on the non-patient-relevant outcome of AHI were reported by 17 studies, and on the non-patient-relevant outcome of ODI, by 4 studies.

Table 2: Matrix of patient-relevant outcomes in research question 1

Study	Outcomes														
	Mortality	Morbidity												QoL	
	All-cause mortality / overall survival	Sleep quality	Excessive daytime sleepiness	Cognitive performance (vigilance)	Cognitive performance (executive functions)	Depressive symptoms	Anxiety symptoms	Psychological symptoms	Fatigue	Somatic symptoms (headache)	Cardiovascular morbidity	SAE	Discontinuation due to AEs	Health-related quality of life	Activities of daily living as well as participation in professional and social life (FOSQ)
Aarab 2010	-	-	●	-	-	-	-	-	-	-	-	○	○	-	-
Aarab 2011	-	○	-	-	-	-	-	●	-	-	-	○	○	○	-
Andrén 2013	-	-	●	-	-	-	-	-	-	-	-	-	-	-	-
Barnes 2004	-	○	●	●	●	●	-	●	-	-	-	○	○	●	●
Bloch 2000	-	-	●	-	-	-	-	-	-	-	-	○	-	-	●
Dal-Fabbro 2014	-	-	●	-	-	-	-	-	-	-	-	-	○	-	-
Durán-Cantolla 2015	-	●	●	-	-	-	-	●	-	-	-	●	-	-	-
Gagnadoux 2017	-	-	●	-	-	-	-	-	●	-	-	○	○	-	-
Godoy 2017	-	●	●	●	-	●	●	-	●	-	-	○	●	-	●
Gotsopoulous 2002	-	○	●	-	●	●	-	●	-	-	-	○	-	-	-
Johnston 2002	-	●	●	-	-	-	-	-	-	-	-	○	○	-	-
Lam 2007	-	-	●	-	-	-	-	-	-	-	-	○	●	●	-
Marklund 2015	-	○	●	●	-	-	-	-	○	●	-	○	○	●	●
Mehta 2001	-	-	-	-	-	-	-	-	-	-	-	○	●	-	-
Petri 2008	-	-	●	-	-	-	-	-	-	-	-	○	●	●	-
Sjöholm 1994[a]	-	-	-	-	-	-	-	-	-	-	-	○	-	-	-
Teixeira 2013[b]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

(continued)

Table 2: Matrix of patient-relevant outcomes in research question 1 (continued)

Study	Outcomes														
	Mortality	Morbidity												QoL	
	All-cause mortality / overall survival	Sleep quality	Excessive daytime sleepiness	Cognitive performance (vigilance)	Cognitive performance (executive functions)	Depressive symptoms	Anxiety symptoms	Psychological symptoms	Fatigue	Somatic symptoms (headache)	Cardiovascular morbidity	SAE	Discontinuation due to AEs	Health-related quality of life	Activities of daily living as well as participation in professional and social life (FOSQ)
TOMADO	-	-	●	-	-	-	-	-	-	-	-	●	●	●	●

●: Data were reported and were usable.

○: Data were reported but not usable for the benefit assessment.

-: No data were reported (no further information), or the outcome was not surveyed.

a: Usable data were available only on the non-patient-relevant outcome of ODI.
b: Usable data were available only on the non-patient-relevant outcome of AHI.

AE: adverse event; FOSQ: Functional Outcomes of Sleep Questionnaire; QoL: health-related quality of life; SAE: serious adverse event

4.2.3 Assessment of the risk of bias at study and outcome levels

The Andrén 2013, Bloch 2000, Gagnadoux 2017, Godoy 2017, Gotsopoulous 2002, and Lam 2007 studies [29,43,45,47,50,53] were rated as having a high risk of bias because it was unclear whether allocation was concealed. For Aarab 2010, Dal-Fabbro 2014, Johnston 2002, and Mehta 2001 [26,38,56,61], it was unclear not only whether allocation was concealed, but also whether the randomization sequence was generated adequately. These 4 studies were therefore rated as having a high risk of bias as well.

The Durán-Cantolla 2015 study [46] was rated as having a high risk of bias because the patient flow from randomization onward was unclear, as was the potential presence of reporting bias.

On the outcome level, further biasing aspects potentially existed as well.

Due to the high risk of bias at study level, the risk of bias for all outcomes was rated as high for these 11 studies.

The Aarab 2011 [36,39-41], Barnes 2004 [37], Marklund 2015 [57,58], Petri 2008 [62], and TOMADO [66,67] studies were rated as having a low risk of bias.

The risk of bias of results on excessive daytime sleepiness was rated as high for the Petri 2008 study [62] because the ITT principle was inadequately implemented. In addition, the outcome assessors were not blinded (patient-reported outcome) in the comparison of MAD versus no treatment.

Due to inadequate implementation of the ITT principle, the risk of bias of results on health-related quality of life was rated as high for the Petri 2008 and TOMADO studies [66,67]. For TOMADO and for the comparison of MAD versus no treatment in Petri 2008, there was also no blinding of outcome assessors (patient-reported outcome). The risk of bias of results on activities of daily living and participation in professional and social life was rated as high for the TOMADO study due to lack of blinding (patient-reported outcome) and inadequate implementation of the ITT principle.

The risk of bias of results on the remaining patient-reported outcomes in the TOMADO study as well as the results on patient-reported outcomes in the Barnes 2004 and Marklund 2015 studies was rated as high due to absent or unclear blinding of outcome assessors (patient-reported outcome).

The risk of bias of results on psychological symptoms from the Aarab 2011 study [36,39-41] was rated as high due to inadequate implementation of the ITT principle.

4.2.4 Results on patient-relevant outcomes

4.2.4.1 Results on all-cause mortality or overall survival

No results were available on all-cause mortality or overall survival. For this outcome, it was therefore not possible to derive a hint of benefit or harm of MAD therapy in comparison with no treatment or placebo treatment.

4.2.4.2 Results on excessive daytime sleepiness

Results on the Epworth Sleepiness Scale (ESS)

In the comparison of MAD therapy versus no treatment or placebo treatment, 14 studies provided data of moderate qualitative certainty of results on the Epworth Sleepiness Scale (ESS). In the evidence syntheses below, the results of the mean differences found in the ESS were standardized and subsequently included in the respective metaanalyses in the form of Hedges' g. (For all metaanalyses on this outcome in research question 1, this was done on the basis of a model with random effects as per Knapp and Hartung.)

Although the Andrén 2013 study reported, for ESS, a statistically significant effect in favour of MAD therapy in comparison with placebo treatment ($p < 0.006$), it was excluded from this study in the qualitative summary because the values at follow-up were missing the measures of dispersion, with no suitable replacement being possible. Some of the remaining 13 studies disregarded data dependency due to crossover design.

Note

To include an approximation of this data dependency in crossover studies, a correlation assumption had to be made (correlation coefficient for the measurements under both treatments). For readability purposes, the shorter term "correlation" will be used instead of the more precise "correlation coefficient" throughout this report. For the outcome of excessive daytime sleepiness as measured by ESS, it was possible to calculate study-specific correlations from 4 studies (2 studies for each research question) since data on the standard error of the mean difference were available or could be calculated from the widths of the confidence interval. For the TOMADO and Johnston 2002 studies (research question 1) as well as for Banhiran 2018 and Phillips 2013 (research question 2), correlations of 0.54 and 0.69 as well as 0.57 and 0.75, respectively, were calculated, thus resulting in a mean correlation of about 0.6. In the main analysis, standard errors were therefore calculated by assuming a correlation of 0.6 for studies which were performed using a crossover design but not analysed accordingly. This applies to all of the outcomes and analyses listed in this report unless a different correlation assumption is explicitly described.

Considerable heterogeneity made a metaanalysis inappropriate for excessive daytime sleepiness as measured using the ESS ($p < 0.001$; $I^2 = 72.0\%$). Therefore, the results were qualitatively summarized. No effects in the same direction were observed because only 4 out of 13 studies showed a statistically significant and clinically relevant effect (see General Methods 5.0

[105,p.184]), because the total weight of these studies was below 50%, and because the prediction interval covered the zero effect. The main analysis for ESS revealed no hint of any benefit of MAD therapy in comparison with no treatment or placebo treatment.

Sensitivity analyses for ESS

Sensitivity analyses were performed to test for potential effects of disregarding the dependence structure using various assumptions about the correlation of observations. To cover a plausible value range of the likely correlation with regard to the dependence structure in studies with a crossover design, 2 further example correlations – 0.4 and 0.8 – were chosen for correction in crossover studies where data dependency was disregarded. Under either correlation assumption, the qualitative evaluation revealed no differences to the main analysis; the total weight of studies with statistically significant and clinically relevant effect remained below 50%, and the prediction interval continued to cover the zero effect. Consequently, no effects in the same direction were found in the sensitivity analyses, either.

No other results on excessive daytime sleepiness were available. For excessive daytime sleepiness, it was therefore not possible to derive a hint of benefit or harm of MAD therapy compared to no treatment or placebo treatment.

Subgroup analyses on ESS

Subgroup analysis by device type

Boil-and-bite appliance versus custom-made appliance

For the included studies, the device type (boil-and-bite appliance versus custom-made appliance) was tested for interaction as a potential effect modifier with regard to excessive daytime sleepiness (ESS).

The test for interaction by device type revealed no statistical significance (p = 0.398).

Twinblock appliance versus monoblock appliance

For the included studies, device design (twinblock appliance versus monoblock appliance) was tested for interaction as a potential effect modifier with regard to excessive daytime sleepiness (ESS). For this purpose, devices with secondary blocking of the maxillary and mandibular parts were considered twinblock appliances even if fixation of the device parts prevents mandibular motion during wear. Only devices with an activator design were considered monoblock appliances in this subgroup analysis.

The test for interaction by device design revealed no statistical significance (p = 0.764).

Devices with individually adjusted degree of protrusion versus devices with preset protrusion

For the included studies, a test for interaction was performed to check the set degree of protrusion (fixed, preset degree of protrusion versus individual adjustment of the degree of protrusion) as a potential effect modifier with regard to excessive daytime sleepiness (ESS).

The test for interaction by protrusion setting revealed no statistical significance (p = 0.351).

Subgroup analysis by severity of OSA

For the included studies, a test for interaction was performed to check the severity of OSA as a potential effect modifier with regard to excessive daytime sleepiness (ESS).

However, the evidence available regarding the severity of OSA did not allow making the planned distinction between studies on patients with mild OSA versus those with moderate OSA; instead, it only allowed distinguishing between studies on patients with mild to moderate OSA versus studies on patients with severe OSA versus studies on patients with mixed or unclear severities of OSA.

The test for interaction by severity of OSA revealed no statistical significance (p = 0.943).

Subgroup analysis by tolerance to PAP therapy

For the included studies, a test for interaction was performed to check patient tolerance to PAP therapy as a potential effect modifier with regard to excessive daytime sleepiness (ESS).

The test for interaction by PAP tolerance revealed no statistical significance (p = 0.106).

Subgroup analysis by comparator intervention

For excessive daytime sleepiness (ESS), a test for interaction was performed to check the intervention in the control group as a potential effect modifier. For that purpose, distinctions were made among:

- placebo treatment with a device (e.g. exclusively the mandibular part or the maxillary part of a twinblock MAD or anti-bruxism device) (placebo device),
- no treatment or placebo treatment not affecting the mandibular position (i.e. without expected therapeutic effect), and
- placebo treatment with a MAD without protrusion (inactive device).

This differentiation is appropriate because placebo devices which lead to an increased vertical dimension [106] or muscle relaxation may indeed have some effect and because the forced jaw closure resulting from an inactive MAD may have some effect as well [106].

The test for interaction by comparator intervention revealed statistical significance (p = 0.032).

As was done in the metaanalysis of all studies, this subgroup analysis assumed a correlation of 0.6 for the main analysis, and 2 sensitivity analyses were then performed using a correlation of 0.4 and 0.8, respectively. The test for interaction by comparator intervention was statistically significant at both correlation coefficients, 0.4 and 0.8 (p = 0.029 and p = 0.035, respectively).

Subgroups MAD versus placebo device and MAD versus inactive device

For the subgroup of MAD versus placebo device, the main analysis assumed a correlation of 0.6, and due to considerable heterogeneity ($p = 0.044$, $I^2 = 62.9\%$), the results were qualitatively summarized. Since only 1 study (Gotsopoulos 2002) provided a statistically significant and clinically relevant result and the weight of this study was below 50%, no effects in the same direction were found. For the sensitivity analyses, at a correlation of 0.4 (meaningful pooling of results possible), no statistically significant effect was found (Hedges' g: −0.31; 95% CI: [−0.66; 0.04]; p = 0.068). For a correlation of 0.8, as for a correlation of 0.6, a qualitative summary of results was generated due to considerable heterogeneity ($p = 0.003$, $I^2 = 78.8\%$). Since only 1 study (Gotsopoulos 2002) provided a statistically significant and clinically relevant result and the weight of this study was below 50%, no effects in the same direction were found.

For the subgroup of MAD versus inactive device, the main analysis under the assumption of a correlation of 0.6 revealed no statistically significant effect (Hedges' g: 0.09; 95% CI: [−0.69; 0.87]; p = 0.669). For the sensitivity analyses, no statistically significant effects resulted at a correlation coefficient of 0.4 (Hedges' g: 0.11; 95% CI: [−0.68; 0.89]; p = 0.617) or 0.8 (Hedges' g: 0.07; 95% CI: [−0.71; 0.85]; p = 0.726).

Subgroup of MAD therapy versus no treatment or placebo treatment not affecting mandibular position

To assess the effect of MAD therapy, the comparison of MAD therapy versus no treatment or placebo treatment not affecting mandibular position seemed most meaningful because it was the only placebo condition without suspected treatment effect.

In the main analysis using a correlation of 0.6, the results were qualitatively summarized due to considerable heterogeneity in the comparison of MAD therapy versus no treatment or placebo treatment not affecting the mandibular position ($p = 0.004$, $I^2 = 70.7\%$). The point estimates of all studies in this subgroup showed the same effect direction in favour of MAD therapy. The proportion of studies showing a statistically significant and clinically relevant effect equalled 51.2% of the total weight. The only study with an estimate > −0.2 had a weight of 13.4%. Hence, effects in the same direction can be assumed.

Afterwards, 2 sensitivity analyses were performed for this subgroup; for crossover studies disregarding dependence structures, the standard error was calculated using an assumed correlation of 0.4 and 0.8.

Due to considerable heterogeneity ($p = 0.026$, $I^2 = 60.7\%$ at a correlation of 0.4, and $p < 0.001$, $I^2 = 83.0\%$ at a correlation of 0.8), the results were qualitatively summarized in each of these sensitivity analyses. In both sensitivity analyses, the point estimates of all studies showed the same effect direction in favour of MAD therapy. The proportion of studies with statistically significant and simultaneously clinically relevant effect equals 50.4% of the total weight of the studies at a correlation of 0.4, and 52.4% of the total weight at a correlation of 0.8. In each case,

the only study with an estimate > −0.2 had weights of 13.5% and 13.3%, respectively; therefore, effects in the same direction can be assumed for both sensitivity analyses as well.

Two studies of this subgroup (Bloch 2000 and TOMADO) each had 2 study arms which were relevant for the intervention group. However, the mentioned metaanalyses each included data from only 1 intervention arm. To capture all of the evidence on the comparison of MAD therapy versus no treatment or placebo treatment not affecting the mandibular position, data from the 2nd intervention arm should also be included in considerations, however.

Both studies showed statistically significant and simultaneously clinically relevant effects in favour of MAD therapy. These results therefore further support the metaanalyses' results.

Hence, for the outcome of excessive daytime sleepiness, as examined by means of ESS, an indication of benefit of MAD therapy in comparison with no treatment or placebo treatment not affecting the mandibular position was derived.

With regard to other factors, the available evidence did not permit testing for effect modification.

Conclusion on benefit regarding excessive daytime sleepiness

Results on excessive daytime sleepiness were available only as measured by ESS. All things considered, for the outcome of excessive daytime sleepiness, it was possible to derive an indication of benefit of MAD therapy only in comparison with no treatment or placebo treatment not affecting the mandibular position.

4.2.4.3 Results on sleep quality

For the comparison of MAD therapy versus no treatment or placebo treatment, usable data of moderate qualitative certainty of results on sleep quality were available from 3 studies: Godoy 2017 collected these data using the Pittsburgh Sleep Quality Index (PSQI) Component 1 (subjective sleep quality) and Component 5 (sleep disturbance); Johnston 2002 used a single question on fatigue upon awakening, and Durán-Cantolla 2015 a single question on any improvement of sleep quality by MAD therapy. The results at the time points 1.5 years (Godoy 2017), 4 to 6 weeks (Johnston 2002), and 12 weeks (Durán-Cantolla 2015) were presented. Due to the differences in survey methods, no metaanalysis of results was compiled.

There was no statistically significant effect for PSQI Component 1 (mean difference [MD]: −0.20; 95% CI: [−1.77; 1.37]; p = 0.796) or for PSQI Component 5 (MD: 0.1; 95% CI: [−1.02; 1.22]; p = 0.857) or for fatigue upon awakening (MD: 0.58; 95% CI: [−0.11; 1.27]; p = 0.094). For the question on improved sleep quality, a statistically significant and clinically relevant effect in favour of MAD therapy was found in comparison with no treatment or placebo treatment (OR: 3.53; 95% CI: [1.37; 9.10]; p = 0.008).

Conclusion on benefit regarding the outcome of sleep quality

All instruments were weighted equally for deriving any benefit. Regarding the question on improved sleep quality, only one single question from a study of moderate qualitative certainty of results (Durán-Cantolla 2015) revealed a statistically significant and clinically relevant effect in favour of MAD therapy in comparison with no treatment or placebo treatment. However, the results of 3 other questions from 2 different studies show no statistically significant effect. Since the only observed effect was found at the time point of 12 weeks, while no effect was observed at an earlier or later time point, the effects cannot be assumed to be time-dependent (as in strictly short-term or long-term effects). With that in mind and in light of the fact that a statistically significant and clinically relevant effect arose in only 1 out of 4 operationalizations, the overall analysis of sleep quality revealed no hint of benefit or harm of MAD therapy in comparison with no treatment or placebo treatment.

4.2.4.4 Results on cognitive performance (vigilance)

Regarding cognitive performance (vigilance), Psychomotor Vigilance Test (PVT) data on errors of omission and false starts were available from 2 studies (Barnes 2004 and Godoy 2017), PVT reaction time data, from 1 study (Godoy 2017), and Multiple Unprepared Reaction Time (MURT) data, from 1 study (Marklund 2015). The various analysis methods (errors of omission, false starts, and reaction time) were analysed separately rather than in a metaanalysis.

Results on the Psychomotor Vigilance Test – errors of omission

In the comparison of MAD therapy versus no treatment or placebo treatment, 1 study (Barnes 2004) provided data of high qualitative certainty of results, and 1 study (Godoy 2017), data of moderate qualitative certainty of results regarding cognitive performance as measured by PVT errors of omission.

Given that the effect in favour of MAD was statistically significant, but not clinically relevant, it was not possible to draw any conclusions from the study of high qualitative certainty of results, Barnes 2004.

The study of moderate qualitative certainty of results, Godoy 2017, was then added, and the results from both studies were pooled. In this case, a metaanalysis was performed based on a model with fixed effect. In the Godoy 2017 study, 5 measurements were taken over the course of the day. Since it was not possible to define 1 of the 5 measurements as the most meaningful, and the values monotonously changed across the 5 measurements for several variables, the first and last measurements of the day were included in the analysis for the purposes of the benefit assessment.

The combined analysis of both studies together with the first measurement from Godoy 2017 produced a non-directed result since the pooled result in favour of MAD therapy was statistically significant but not clinically relevant (Hedges' g: −0.17; 95% CI: [−0.33; −0.01]; p = 0.041). The combined analysis of both studies together with the last measurement from

Godoy 2017 produced a non-directed result as well because the pooled result in favour of MAD therapy was also statistically significant but not clinically relevant (Hedges' g: −0.17; 95% CI: [−0.33; −0.01]; p = 0.043).

Results on the Psychomotor Vigilance Test – false starts

In the comparison of MAD therapy versus no treatment or placebo treatment, 1 study (Barnes 2004) provided data of high qualitative certainty of results and 1 further study (Godoy 2017) provided data of moderate qualitative certainty of results to assess cognitive performance (vigilance) as measured by means of the PVT analysis of false starts.

Due to the absence of any statistically significant effect, it was not possible to draw any conclusions from the study of high qualitative certainty of results, Barnes 2004.

As was done for the results of the Psychomotor Vigilance Test (errors of omission), a metaanalysis of the results from Barnes 2004 and Godoy 2017 was then compiled (model with fixed effect, once using the first and once using the last measurement of the day from Godoy 2017). The combined analysis of both studies together with the first measurement from Godoy 2017 produced a non-directed result since the pooled result was not statistically significant (Hedges' g: −0.03; 95% CI: [−0.19; 0.13]; p = 0.703). The combined analysis of both studies together with the last measurement from Godoy 2017 revealed a non-directed result as well because the pooled result was also not statistically significant (Hedges' g: −0.03; 95% CI: [−0.19; 0.13]; p = 0.716).

Results on reaction time

For the comparison of MAD therapy versus no treatment or placebo treatment regarding the "reaction time" aspect of the outcome of cognitive performance (vigilance), 1 study (Marklund 2015) provided data of high qualitative certainty of results from the MURT test, and 1 study (Godoy 2017) provided data of moderate qualitative certainty of results from PVT.

In the absence of a statistically significant effect, it was not possible to draw any conclusions from the study of high qualitative certainty of results, Marklund 2015.

As was done for the Psychomotor Vigilance Test (errors of omission), a metaanalysis of the results from Marklund 2015 and Godoy 2017 was then compiled (model with fixed effect, once using the first and once using the last measurement of the day from Godoy 2017). The combined analysis of both studies together with the first measurement from Godoy 2017 produced a non-directed result since the pooled result was not statistically significant (Hedges' g: −0.12; 95% CI: [−0.48; 0.24]; p = 0.518). The combined analysis of both studies together with the last measurement from Godoy 2017 revealed a non-directed result as well because the pooled result was also not statistically significant (Hedges' g: −0.12; 95% CI: [−0.48; 0.23]; p = 0.501).

Conclusion on benefit regarding the outcome of cognitive performance (vigilance)

All instruments and all evaluation types (errors of omission, false starts, and reaction time) were weighted equally for deriving any benefit. Since none of the instruments revealed any statistically significant and simultaneously clinically relevant effect, it was not possible to derive an overall hint of benefit or harm of MAD therapy in comparison with no treatment or placebo treatment for cognitive performance (vigilance).

4.2.4.5 Results on cognitive performance (executive functions)

For cognitive performance (executive functions), 2 studies (Barnes 2004 and Gotsopoulos 2002) provided data on the digit span backwards, the Trail Making Test B, and the F-A-S test (proper name of the word fluency test) or Controlled Oral Word Association Test (COWAT). In addition, 1 study (Barnes 2004) provided data on 2 procedure versions of the Paced Auditory Serial Addition Task (PASAT). Due to the difference in survey methods, no metaanalysis of the results of different instruments was compiled.

Results on the Trail Making Test B

In the comparison of MAD therapy versus no treatment or placebo treatment regarding cognitive performance (executive functions) in the form of the analysis of the Trail Making Test B, 1 study (Barnes 2004) provided data of high qualitative certainty of results, and 1 study (Gotsopoulos 2002) provided data of moderate qualitative certainty of results.

Due to the absence of any statistically significant effect, it was not possible to draw any conclusions from the study of high qualitative certainty of results, Barnes 2004.

The study with moderate qualitative certainty of results, Gotsopoulos 2002, was then added, and the results from both studies were considered jointly. In this case, a metaanalysis was performed based on a model with fixed effect.

The combined analysis of both studies revealed a non-directed result because the pooled result was not statistically significant (Hedges' g: 0.01; 95% CI: [−0.11; 0.14]; p = 0.821).

Results on the Paced Auditory Serial Addition Task (PASAT)

In the comparison of MAD therapy versus no treatment or placebo treatment, 1 study (Barnes 2004) provided data of high qualitative certainty of results for 2 procedure versions (interstimulus interval of 1.2 and 2.4 seconds) for the PASAT analysis of cognitive performance (executive functions).

The separate analysis of Barnes 2004 for PASAT with an interstimulus interval of 1.2 seconds showed a statistically significant and clinically relevant effect in favour of MAD therapy (MD: −0.80; 95% CI: [−1.01; −0.59]; p < 0.001; Hedges' g: −1.01; 95% CI: [−1.33; −0.69]).

The separate analysis of Barnes 2004 for PASAT with an interstimulus interval of 2.4 seconds showed no statistically significant effect (MD: 0.00; 95% CI: [−0.28; 0.28]; $p > 0.999$; Hedges' g: 0.00; 95% CI: [−0.26; 0.26]).

Results on digit span backward

In the comparison of MAD therapy versus no treatment or placebo treatment, 1 study (Barnes 2004) provided data of high qualitative certainty of results and 1 study (Gotsopoulos 2002), data of moderate qualitative certainty of results regarding cognitive performance (executive functions) using an analysis of digit span backward.

Given that the effect in favour of MAD was statistically significant, but not clinically relevant, it was not possible to draw any conclusions from the study of high qualitative certainty of results, Barnes 2004.

The study with moderate qualitative certainty of results, Gotsopoulos 2002, was then added, and the results from both studies were considered jointly. In this case, a metaanalysis was performed based on a model with fixed effect.

Joint consideration of the two studies revealed a non-directed result: due to meaningful heterogeneity ($p = 0.031$; $I^2 = 78.4\%$), pooling of results was not meaningful, and the results were qualitatively summarized. No studies with a statistically significant and simultaneously clinically relevant effect were available; consequently, it was not possible to derive any effects in the same direction.

Results on the F-A-S test / Controlled Oral Word Association Test (COWAT)

In the comparison of MAD therapy versus no treatment or placebo treatment, 1 study (Barnes 2004) provided data of high qualitative certainty of results and 1 study (Gotsopoulos 2002) provided data of moderate qualitative certainty of results regarding cognitive performance (executive functions) using the analysis of the F-A-S test / COWAT.

Due to the absence of any statistically significant effect, it was not possible to draw any conclusions from the study of high qualitative certainty of results, Barnes 2004.

The study with moderate qualitative certainty of results, Gotsopoulos 2002, was then added, and the results from both studies were considered jointly. In this case, a metaanalysis was performed based on a model with fixed effect.

The combined analysis of both studies revealed a non-directed result because the pooled result was not statistically significant (Hedges' g: −0.04; 95% CI: [−0.16; 0.09]; $p = 0.592$).

Conclusion on benefit regarding the outcome of cognitive performance (executive functions)

All instruments used to measure cognitive performance (executive functions) were weighted equally when drawing the conclusion on benefit. This revealed a statistically significant and

clinically relevant effect of MAD therapy in comparison with no treatment or placebo treatment only for PASAT using an interstimulus interval of 1.2 seconds on the basis of a study of high qualitative certainty of results (Barnes 2004); hence, there is an indication of benefit of MAD therapy in comparison with no treatment or placebo treatment for this instrument. This result is based on data from one study on the subgroup of MAD versus placebo tablet.

The results for the other operationalization of the PASAT or other instruments revealed effects which were not statistically significant effect or were statistically significant, but not clinically relevant. Since a statistically significant and clinically relevant effect was found in only 1 out of 5 operationalizations or instruments, the overall analysis of cognitive performance (executive functions) resulted in no hint of benefit or harm of MAD therapy in comparison with no treatment or placebo treatment.

4.2.4.6 Results on depressive symptoms

In the comparison of MAD therapy versus no treatment or placebo treatment, 3 studies of moderate qualitative certainty of results (Barnes 2004, Godoy 2017, Gotsopoulos 2002) provided data on depressive symptoms as measured by the Beck Depression Inventory (BDI).

A metaanalysis was conducted of the results from all 3 studies. Since the studies differed considerably in design and research question, particularly regarding the intervention, control, and population, an analysis was done based on a model with random effects as per Knapp and Hartung.

No statistically significant effect was found (Hedges' g: −0.16; 95% CI: [−0.44; 0.11]; p = 0.126).

A qualitative evidence synthesis did not produce any different results because there were no effects in the same direction.

Conclusion on benefit regarding the outcome of depressive symptoms

For the BDI, no statistically significant effect was found. No results were available on any other instruments.

Overall, for depressive symptoms, there was therefore no hint of benefit or harm of MAD therapy in comparison with no treatment or placebo treatment.

4.2.4.7 Results on anxiety symptoms

In the comparison of MAD therapy versus no treatment or placebo treatment, 1 study (Godoy 2017) provided data of moderate qualitative certainty of results regarding anxiety symptoms as measured by the Beck Anxiety Inventory (BAI).

No statistically significant effect was found (MD: −9.6; 95% CI: [−26.3; 7.1]; p = 0.249).

Conclusion on benefit regarding the outcome of anxiety symptoms

Regarding BAI, no statistically significant effect was found. No results were available on any other instruments.

Overall, for anxiety symptoms, there was therefore no hint of any benefit or harm of MAD therapy in comparison with no treatment or placebo treatment.

4.2.4.8 Results on psychological symptoms

Regarding psychological symptoms, 1 study (Aarab 2011) provided data on the Symptom Checklist-90-Revised (SCL-90-R), 2 studies (Barnes 2004 and Gotsopoulos 2002) on the Profile of Mood States (POMS), and 1 study (Durán-Cantolla 2015) on 1 single question each on well-being and mood. Due to the difference in survey methods, no metaanalysis was compiled of the POMS and SCL-90-R instruments as well as of the single questions on well-being and mood from Durán-Cantolla 2015.

Results on the Profile of Mood States

In the comparison of MAD therapy versus no treatment or placebo treatment, 2 studies of moderate qualitative certainty of results (Barnes 2004, Gotsopoulos 2002) provided data on psychological symptoms as measured by POMS. The total score was analysed.

A metaanalysis of the results of both studies was compiled. In this case, an analysis was done based on a model with fixed effect.

No statistically significant effect was found (Hedges' g: −0.09; 95% CI: [−0.22; 0.03]; p = 0.152).

Results on the Symptom Checklist-90-Revised

In the comparison of MAD therapy versus no treatment or placebo treatment, 1 study (Aarab 2011) provided data of moderate qualitative certainty of results on the outcome of psychological symptoms as measured by SCL-90-R. The total score (mean psychological distress) was analysed.

No statistically significant effect was found (MD: 8.8; 95% CI: [−12.03; 29.63]; p = 0.397).

Results on well-being (single question)

In the comparison of MAD therapy versus no treatment or placebo treatment, 1 study (Durán-Cantolla 2015) provided data of moderate qualitative certainty of results regarding psychological symptoms from a single question about well-being.

No statistically significant effect was found (OR: 2.1; 95% CI: [0.82; 5.40]; p = 0.126).

Results on mood (single question)

For the comparison of MAD therapy versus no treatment or placebo treatment, 1 study (Durán-Cantolla 2015) provided data of moderate qualitative certainty of results regarding psychological symptoms as measured by a single question on mood.

No statistically significant effect was found (OR: 1.14; 95% CI: [0.37; 3.54]; p = 0.864).

Conclusion on benefit regarding the outcome of psychological symptoms

The results on the two single questions were weighted less than the other instruments in the benefit assessment. POMS and SCL-90-R were weighted equally in the benefit assessment, and only the total scores of the instruments were taken into account. Since none of the instruments revealed any statistically significant effect, it was not possible to derive any overall hint of benefit or harm of MAD therapy in comparison with no treatment or placebo treatment for psychological symptoms.

4.2.4.9 Results on somatic symptoms – headache

In the comparison of MAD therapy versus no treatment or placebo treatment, 1 study (Marklund 2015) provided data of moderate qualitative certainty of results from 4 different operationalizations of "somatic symptoms – headache." Due to the data distribution, representation, and scale, calculating a mean difference did not seem meaningful for 3 out of the 4 operationalizations (days with a headache, longest headache, and headache intensity).

The p-values reported in the study (p = 0.87 for days with a headache; p = 0.56 for longest headache; p = 0.54 for headache intensity) showed no statistically significant effect for any of the operationalizations.

For the 4th operationalization, "headache present", no statistically significant effect was found (OR: 1.08; 95% CI: [0.44; 2.65]; p = 0.93). No further results on the outcome of "somatic symptoms – headache" were available.

Conclusion on benefit regarding the outcome "somatic symptoms – headache"

For "somatic symptoms – headache", there was therefore no hint of benefit or harm of MAD therapy in comparison with no treatment or placebo treatment.

4.2.4.10 Results on cardiovascular morbidity

No results were available on cardiovascular morbidity. For this outcome, it was therefore not possible to derive a hint of benefit or harm of MAD therapy in comparison with no treatment or placebo treatment.

4.2.4.11 Results on fatigue

Regarding fatigue, 1 study (Gagnadoux 2017) provided data from a single question, and 1 study (Godoy 2017), from the Modified Fatigue Impact Scale (MFIS). Due to the differences in survey methods, no metaanalysis of results was compiled.

Results on fatigue (single question)

On fatigue, 1 study (Gagnadoux 2017) provided usable data of moderate qualitative certainty of results from a single question. In the study, a MAD was compared with a placebo device (maxillary portion of the device). In the analysis, the "fatigue considerably reduced" category was compared with the other categories because considerable reduction was seen as sufficiently patient-relevant in view of the research question. A statistically significant and clinically relevant effect was found (OR: 8.14; 95% CI: [2.29; 28.90]; $p < 0.001$) for the comparison of MAD versus placebo device.

Results on fatigue – Modified Fatigue Impact Scale

In the comparison of MAD therapy versus no treatment or placebo treatment, 1 study (Godoy 2017) provided usable data of moderate qualitative certainty of results for fatigue as measured by the Modified Fatigue Impact Scale (MFIS).

No statistically significant effect was found (MD: −12.7; 95% CI: [−39.07; 13.67]; $p = 0.332$).

Conclusion on benefit regarding the outcome of fatigue

For the single question on fatigue, a statistically significant and clinically relevant effect was found in favour of MAD in comparison with a placebo device on the basis of a study of moderate qualitative certainty of results with a follow-up period of 2 months (Gagnadoux 2017).

For MFIS, no statistically significant effect was found on the basis of the Godoy 2017 study with a follow-up period of 18 months. No further results on fatigue were available. The results on the single question and the results on MFIS were weighted equally in the benefit assessment. The overall analysis revealed a hint of benefit of MADs in comparison with placebo devices regarding fatigue because the results on the single question are based on a much larger study, Gagnadoux 2017 with 150 patients, than the MFIS results, which are based on Godoy 2017 with 30 patients; therefore, the results stem from a larger population with equal certainty of results, and the result from Gagnadoux 2017 represented a large effect.

4.2.4.12 Results on health-related quality of life

Results of the 36-Item Short Form Health Survey (SF-36)

In the comparison of MAD therapy versus no treatment or placebo treatment regarding health-related quality of life, 5 studies (Barnes 2004, Lam 2007, Marklund 2015, Petri 2008, TOMADO) provided usable data of moderate qualitative certainty of results on the 36-Item Short Form Health Survey (SF-36).

For each of the SF-36 summary scores (SF-36 physical health summary score and SF-36 mental health summary score), data were available from 3 studies. In order to take into account the results from the other studies as well, a metaanalysis of SF-36 was conducted for the following subscores: SF-36 general health, SF-36 vitality, SF-36 physical functioning, SF-36 bodily pain, SF-36 social role functioning, SF-36 mental health, SF-36 physical role functioning, and SF-36 emotional role functioning.

Since for all examined scales and subscales, the studies (considerably) differed in design and research question, particularly regarding the intervention, control, and population, a metaanalysis was compiled for 3 or 4 studies based on a model with random effects as per Knapp and Hartung.

No statistically significant and simultaneously clinically relevant effect was found for any of the summary scores or any of the subscales.

Qualitative evidence synthesis did not reveal any other results for any of the subscales since there were no effects in the same direction.

Conclusion on benefit regarding the outcome of health-related quality of life

For SF-36, both the summary scores and the subscores were used to derive any benefit. No statistically significant and simultaneously clinically relevant effect was found for any of the summary scores or subscores. No results were available on any other instruments.

For health-related quality of life, it was therefore not possible to derive a hint of benefit or harm of MAD therapy when compared to no treatment or placebo treatment.

4.2.4.13 Results on activities of daily living as well as participation in professional and social life

For activities of daily living as well as participation in professional and social life, 4 studies (Barnes 2004, Godoy 2017, Marklund 2015, and TOMADO) provided data on the Functional Outcomes of Sleep Questionnaire (FOSQ). Additionally, 1 study (Bloch 2000) provided data on single questions about activities of daily living. Due to the differences in survey methods, no metaanalysis of results was compiled.

Results on the Functional Outcomes of Sleep Questionnaire (FOSQ)

In the comparison of MAD therapy versus no treatment or placebo treatment, 4 studies (Barnes 2004, Godoy 2017, Marklund 2015, TOMADO) provided data of moderate qualitative certainty of results on activities of daily living and participation in professional and social life as measured by the Functional Outcomes of Sleep Questionnaire (FOSQ).

Due to substantial heterogeneity ($p = 0.003$; $I^2 = 78.2\%$), the results were qualitatively summarized. Since the studies differed considerably in design and research question, particularly regarding the intervention, control, and population, an analysis was done based on

a model with random effects as per Knapp and Hartung. The prediction interval covered the zero effect, and only 1 study had a statistically significant result. Therefore, there were no effects in the same direction.

Consequently, for activities of daily living and participation in professional and social life as measured by FOSQ, it was not possible to derive a hint of benefit or harm of MAD therapy in comparison with no treatment or placebo treatment.

Results from single questions on activities of daily living

In the comparison of MAD therapy versus no treatment or placebo treatment, 1 study (Bloch 2000) provided data of moderate qualitative certainty of results regarding activities of daily living and participation in professional and social life from single questions regarding negative impact on daily tasks, such as energy level and performance; these data were analysed for 2 different custom-made devices. The analysis and interpretation of study results were difficult since the exact methodology and the clinical relevance threshold remained unclear. Furthermore, due to the distribution found, calculating the mean and standard deviation was not meaningful. According to the study publication, however, statistically significant effects were observed for all 3 single questions (MAD 1 Herbst appliance: $p < 0.03$ for all 3 questions; MAD 2 monoblock: $p < 0.001$ for all 3 questions).

Conclusion on benefit regarding activities of daily living as well as participation in professional and social life

The results on the single questions were weighted lower than the FOSQ when deriving any benefit. Only the FOSQ total score was taken into account. Statistically significant effects were reported on the basis of the single questions from 1 study (Bloch 2000), but the qualitative summary of results of the FOSQ showed no effects in the same direction. Since the FOSQ results represent a metaanalysis of 4 studies on a multidimensional instrument, there is overall no hint of benefit or harm of MAD therapy in comparison with no treatment or placebo treatment.

4.2.4.14 Results on serious adverse events and discontinuation due to adverse events

Generally, the studies surveyed or reported outcomes unsystematically. A total of 6 studies (Durán-Cantolla 2015, Godoy 2017, Lam 2007, Mehta 2001, Petri 2008, and TOMADO) provided specific data of moderate qualitative certainty of results on the occurrence of serious adverse events or discontinuation due to adverse events. Additional studies with reported results lacked data, for instance on the number of events, or it was unclear whether a reported discontinuation was due to adverse events or to other causes.

Reported specific adverse events generally included, for instance hypersalivation, temporomandibular joint pain, toothaches, or occlusion impairment. Only in isolated cases did the description of events suggest that they were severe. In fact, studies repeatedly described the AEs as exclusively mild or only transient. The available information on serious adverse events

did not show any higher frequency of serious adverse events in MAD therapy versus no treatment or placebo.

In terms of discontinuation due to adverse events, the 5 studies with usable data reported between 1 and 4 discontinuations, corresponding to 1.2% to 12%, in the intervention groups, compared to 0 to 2 discontinuations, corresponding to 0% to 7%, in the control groups. In the other studies relevant for research question 1, no data were available on this outcome.

Apparently, no harm is discernible on the basis of the available data on the two outcomes of discontinuation due to adverse events and serious adverse events. On the basis of the available data, there is no hint of benefit or harm of MAD therapy in comparison with no treatment or placebo treatment.

4.2.5 Results on AHI and ODI

Since AHI and ODI are neither patient-relevant outcomes nor validated surrogate outcomes, they were not used in the benefit assessment. The results are discussed only as supplementary information in the details of the report.

4.3 Results on research question 2 – MAD versus PAP therapy

4.3.1 Characteristics of the studies included in the assessment

For research question 2, a total of 20 studies were included, of which the results of 17 studies were used in the benefit assessment. The results of 3 studies were excluded from the benefit assessment due to unusable data: These studies involved either an excessive percentage of patients excluded from the analysis (Schütz 2013 [100] and COMET [72,73]) or from an excessive difference between intervention groups in the percentage of patients excluded from the analysis (de Vries 2019 [74]). The information provided below applies to the 17 studies with usable data.

Thirteen studies had 2 study arms comparing MAD versus PAP therapy, while 4 studies had 3 study arms: In addition to the MAD and PAP study arms, 2 of these studies had an arm with an inactive device, 1 study had an arm with a placebo tablet, and 1 other study, an arm with conservative treatment measures, which were administered as accompanying therapy in the MAD and PAP study arms. Four studies used a parallel-group design, and 13 studies, a crossover design. The studies were conducted between 1991 and 2016 (in some cases, the figures reflected the recruitment period, and for 8 studies, the available documents did not provide any information on the study period). Between 20 and 126 patients were randomized, totalling 915 patients across all 17 studies. Patients in crossover studies were entered into the calculation of the patient total only once; patients in 3-arm studies with parallel design were included in the calculation only if they were in the study arms of interest for research question 2.

Custom-made MADs were used in 15 studies, and boil-and-bite appliances, in 2 studies. A total of 11 studies used twinblock appliances, a total of 3 studies, monoblock appliances, and

2 studies, both monoblock and twinblock appliances; in 1 study, no information was provided on appliance type. Among the studies with usable data on ESS, 10 studies used twinblock appliances, 1 study, monoblock appliances, and 2 studies, both monoblock and twinblock appliances. The treatment durations were between 1 month and 2 years.

All studies included patients with OSA. In 16 studies, AHI was among the inclusion criteria. For AHI, 1 study defined merely an upper limit for inclusion, namely AHI < 50: 8 studies defined both a lower AHI limit (AHI > 5 or ≥ 5 in 4 studies, AHI ≥ 10 in 1 study, AHI ≥ 15 in 2 studies, AHI ≥ 20 in 1 study) and an upper AHI limit (AHI ≤ 30 in 2 studies, AHI ≤ 40 in 2 studies, AHI ≤ 45 in 1 study, AHI ≤ 50 in 1 study, AHI ≤ 55 in 1 study, and AHI ≤ 60 in 1 study). For AHI, 7 studies defined only a lower limit as an inclusion criterion (AHI ≥ 5 in 5 studies, AHI > 10 in 1 study, and AHI ≥ 20 in 1 study).

4.3.2 Overview of assessment-relevant outcomes

Data on patient-relevant outcomes were extractable from 15 studies; 2 studies provided usable data only on the non-patient-relevant outcomes of AHI (Glos 2016 [86] and Randerath 2002 [27]) and ODI (Glos 2016). Table 3 presents an overview of the available data on patient-relevant outcomes as provided by the included studies. Two studies reported data on sleep quality, but these data were not usable for the benefit assessment. Another four studies provided unusable data, each on a different outcome, namely mortality, excessive daytime sleepiness, cognitive performance (vigilance), and psychological symptoms. Usable data on the outcomes of sleep quality, excessive daytime sleepiness, cognitive performance (vigilance), and psychological symptoms were, however, available from other studies. None of the studies reported data on the outcomes of somatic symptoms or cardiovascular morbidity.

Fifteen studies provided data on the non-patient-relevant outcome of AHI, and 7 studies, on the non-patient-relevant outcome of ODI.

Table 3: Matrix of patient-relevant outcomes in research question 2

Study	Outcomes													
	Mortality	Morbidity											QoL	
	All-cause mortality / overall survival	Sleep quality	Excessive daytime sleepiness	Cognitive performance (vigilance)	Cognitive performance (executive functions)	Depressive symptoms	Anxiety symptoms	Psychological symptoms	Somatic symptoms	Cardiovascular morbidity	SAE	Discontinuation due to AEs	Health-related quality of life	Activities of daily living as well as participation in professional and social life (FOSQ)
Aarab 2011	-	-	●	-	-	-	-	●	-	-	○	●	-	-
Arya 2014	-	●	●	-	-	-	-	-	-	-	○	-	-	-
Banhiran 2018	-	-	●	-	-	-	-	-	-	-	○	●	-	●
Barnes 2004	-	-	●	●	●	●	-	●	-	-	○	○	●	●
Dal-Fabbro 2014	-	-	●	-	-	-	-	-	-	-	-	○	-	-
El-Solh 2017	○	●	●	-	-	-	-	-	-	-	●	-	●	-
Engleman 2002	-	-	●	●	●	●	●	-	-	-	○	-	●	●
Ferguson 1996	-	○	○	-	-	-	-	-	-	-	○	●	-	-
Ferguson 1997	-	○	●	-	-	-	-	-	-	-	○	●	-	-
Gagnadoux 2009	-	-	●	-	●	-	-	-	-	-	○	-	●	-
Glos 2016	-	-	-	-	-	-	-	-	-	-	-	○	-	-
Hoekema 2006	-	-	●	○	-	●	●	○	-	-	○	○	●	●
Lam 2007	-	-	●	-	-	-	-	-	-	-	○	●	●	-
Phillips 2013	-	-	●	●	-	-	-	-	-	-	-	-	●	●
Randerath 2002	-	-	-	-	-	-	-	-	-	-	○	-	-	-
Tan 2002	-	-	●	-	-	-	-	-	-	-	○	●	-	-
Yamamoto 2019	-	-	●	-	-	-	-	-	-	-	-	-	-	-

(continued)

Table 3: Matrix of patient-relevant outcomes in research question 2 (continued)

●: Data were reported and were usable.
○: Data were reported but not usable for the benefit assessment.
-: No data were reported (no further information), or the outcome was not surveyed.
AE: adverse event; FOSQ: Functional Outcomes of Sleep Questionnaire; QoL: health-related quality of life; SAE: serious adverse event

4.3.3 Assessment of the risk of bias at study and outcome levels

Dal-Fabbro 2014, Engleman 2002, Ferguson 1996, Ferguson 1997, Gagnadoux 2009, and Tan 2002 [26,80-83,102] were rated as having a high risk of bias since it remained unclear whether the randomization sequence had been generated adequately. Arya 2014, Lam 2007, Phillips 2013, and Yamamoto 2019 [29,69,97,103] were rated as having a high risk of bias since it remained unclear whether group allocation had been concealed. (This aspect was considered unclear for Dal-Fabbro 2014, Engleman 2002, Ferguson 1996, Ferguson 1997, and Tan 2002 as well.) For Gagnadoux 2009, it was also unclear whether the study suffered from reporting bias.

On the outcome level, further biasing aspects potentially existed as well.

Due to the high risk of bias at study level, the risk of bias for all outcomes was rated as high for these 10 studies.

Aarab 2011 [36,39-41], Banhiran 2018 [28], Barnes 2004 [37], El-Solh 2017 [76], and Hoekema 2006 [30,88-95] were rated as having a low risk of bias.

The outcomes of excessive daytime sleepiness and sleep quality from El-Solh 2017 [76] and the outcomes of excessive daytime sleepiness, activities of daily living as well as participation in professional and social life from Banhiran 2018 [28] were rated as having a high risk of bias due to lack of blinding of the outcome assessors (patient-reported outcome) and inadequate implementation of the ITT principle.

For El-Solh 2017, the risk of bias of the results on health-related quality of life was rated as high because the ITT principle was inadequately implemented.

For Aarab 2011, the risk of bias of the results on excessive daytime sleepiness and psychological symptoms was rated as high due to lack of blinding of the outcome assessors (patient-reported outcome). For the outcome of psychological symptoms and 12-month data on excessive daytime sleepiness, the ITT principle was inadequately implemented as well.

The risk of bias of the results on patient-reported outcomes in the Barnes 2004 and Hoekema 2006 studies was rated as high due to lack of blinding of outcome assessors (patient-reported outcome).

The risk of bias for all remaining outcomes of the studies with a low risk of bias on the study level was rated as low.

4.3.4 Results on patient-relevant outcomes

4.3.4.1 Results on all-cause mortality or overall survival

No usable results were available on all-cause mortality or overall survival. Consequently, for this outcome, it was not possible to derive a hint of greater benefit or harm of MAD therapy compared to PAP therapy.

4.3.4.2 Results on excessive daytime sleepiness

For excessive daytime sleepiness, 13 studies (Aarab 2011, Banhiran 2018, Barnes 2004, Dal-Fabbro 2014, El-Solh 2017, Engleman 2002, Ferguson 1997, Gagnadoux 2009, Hoekema 2006, Lam 2007, Phillips 2013, Tan 2002, Yamamoto 2019) provided data on ESS, and 1 study (Arya 2014), on a single question from the Berlin Questionnaire for sleep apnea (BQ). Due to the difference in survey methods, the single question was not included in the metaanalysis of ESS results.

Results on the Epworth Sleepiness Scale (ESS)

For the comparison of MAD versus PAP therapy, 13 studies (Aarab 2011, Banhiran 2018, Barnes 2004, Dal-Fabbro 2014, El-Solh 2017, Engleman 2002, Ferguson 1997, Gagnadoux 2009, Hoekema 2006, Lam 2007, Phillips 2013, Tan 2002, Yamamoto 2019) provided usable data of moderate qualitative certainty of results on the ESS.

The results were combined in a metaanalysis using the Knapp and Hartung method. For this purpose, the results of the mean differences in the ESS were standardized and then included in the metaanalysis using Hedges' g. As in research question 1, the metaanalysis involved calculating standard errors with an assumed correlation of 0.6 (see Section 4.2.4.2) for studies which were performed using a crossover design but not analysed accordingly.

Due to heterogeneity ($p = 0.034$; $I^2 = 46.3\%$), meaningful pooling the effect estimators of the primary studies was not possible. The results were qualitatively summarized. No effects in the same direction were found because only 1 study exhibited a statistically significant and clinically relevant effect above the irrelevance threshold of 0.2 and the prediction interval covered the zero effect.

Aarab 2011 provided 12-month data in addition to the 6-month data included in the metaanalysis due to the temporal proximity to the other studies and due to the adequately implemented ITT principle. No statistically significant effect was found for these data (MD: −1.2; 95% CI: [−4.79; 2.39]; $p < 0.500$; Hedges' g: −0.26; 95% CI: [−0.94; 0.43]; $p = 0.462$). Hence, it was not possible to derive a hint of greater benefit or harm of MAD therapy compared to PAP therapy on the basis of a superiority test regarding excessive daytime sleepiness as measured by the ESS.

Noninferiority testing of MAD therapy in comparison with PAP therapy regarding ESS

In addition to superiority testing, the investigation called for noninferiority testing of MAD therapy in comparison with PAP therapy. For this purpose, a threshold was to be defined for the main symptom of excessive daytime sleepiness (as measured by ESS) up to which a disadvantage of MAD therapy in comparison with PAP therapy seems tolerable. Simultaneously, no relevant disadvantage of MAD therapy versus PAP therapy must have been found regarding the other outcomes. As part of an initial scoping search, 8 publications [107-114] were found in which an ESS score difference of 2 points was defined as a noninferiority margin (or as an equivalence range in a publication). However, these publications did not provide any data suitable for empirically supporting the adequacy of an ESS score difference of 2 points as a noninferiority margin. Consequently, the Hedges' g standardized mean difference was used in the noninferiority testing. Given that the advantages of MAD therapy over PAP therapy (MADs are typically experienced as less uncomfortable, are better tolerated than PAP therapy [1,34,35], and are associated with higher compliance [115]) and the fact that MAD therapy cannot be assumed to be relevantly inferior to PAP therapy in terms of overall mortality or cardiovascular outcomes (see Section A4.4.4 of the full report), the noninferiority margin of 0.25, which was defined in the report plan, was considered adequate for excessive daytime sleepiness as measured by ESS. Any differences to the disadvantage of MAD therapy which were below this noninferiority margin seemed tolerable.

The main analysis again used an assumed correlation of 0.6 for any crossover studies disregarding data dependency. The results were qualitatively summarized. The prediction interval overlapped the noninferiority margin of 0.25. At an upper CI limit of 0.25 for Yamamoto 2019, the rounding precision in this study determined whether or not the effects were overall in the same direction with regard to the margin of 0.25. After all, the weight of the studies with an upper CI limit < 0.25 was 53.0% when Yamamoto 2019 was included and 44.4% when it was excluded. The total weight of studies with an estimate > 0.25 was 19.5%. Hence, the results were borderline for effects being in the same direction.

Sensitivity analyses for noninferiority testing regarding ESS

To test the potential effects of some crossover studies' disregard of dependence structures under various assumptions about the correlation of observations, sensitivity analyses on the consideration of dependence structures were conducted alongside the main analysis, which assumed a correlation of 0.6. The sensitivity analyses assumed correlations of 0.4 and 0.8 in order to cover a plausible value range for the size of the likely correlation (see Section 4.2.4.2).

In the sensitivity analysis with an assumed correlation of 0.4, it was possible to pool the effect estimates of the primary studies. It found noninferiority at a margin of 0.25 (Hedges' g: 0.07; 95% CI: [−0.06; 0.20]). In the sensitivity analysis with an assumed correlation of 0.8, meaningful pooling of the primary studies' effect estimators was not possible due to heterogeneity ($p = 0.009$, $I^2 = 54.9\%$), and the results were qualitatively summarized. The prediction interval overlapped the noninferiority margin of 0.25. The weight of the studies with

an upper CI limit < 0.25 was 59.3%, while the weight of studies with an estimate > 0.25 was 17.0%. Consequently, effects in the same direction were found at a margin of 0.25.

All things considered, the results of these two metaanalyses therefore do not contradict the results of the main analysis.

However, due to the fact that substantial statistical heterogeneity was found in the main analysis (at a correlation of 0.6) as well as in one of the sensitivity analyses (at a correlation of 0.8) and that this heterogeneity made meaningful pooling of results impossible, 2 further sensitivity analyses each were performed for the main analysis and both sensitivity analyses. In total, 9 analyses were conducted. The additional sensitivity analyses to examine heterogeneity were performed both as a shifting test and as a metaanalysis without the Engleman 2002 study, which was largely responsible for the heterogeneity.

Shifting test

In the shifting test, the point estimates of the 3 studies with the largest effects in favour of MAD therapy (El-Solh 2017, Ferguson 1997, and Gagnadoux 2009) were shifted in the direction of the zero effect until their point estimates were identical with the point estimate of the study with the fourth largest effect in favour of MAD therapy (Phillips 2013). The analyses described below therefore assumed, for these 3 studies, a lesser benefit of MAD therapy than was observed in the actual studies.

The main analysis of the shifting test was performed assuming a correlation of 0.6 for those crossover studies that had disregarded data dependency. The results were homogeneous, and pooling of the effect estimators was therefore possible. With an overall effect estimator of Hedges' g: 0.07; 95% CI: [−0.05; 0.19], noninferiority with a margin of 0.25 was found since the upper CI limit was below 0.25.

The shifting test was combined not only with the assumed correlation of 0.6 (main analysis), but also with the two correlation assumptions of the sensitivity analyses (0.4 and 0.8). The sensitivity analysis assuming a correlation of 0.4 resulted in an overall effect estimate of Hedges' g of 0.08 at a 95% CI of [−0.04; 0.20], revealing noninferiority at a margin of 0.25.

In the sensitivity analysis using an assumed correlation of 0.8, meaningful pooling of the primary studies' effect estimators was not possible due to heterogeneity ($p = 0.033$; $I^2 = 46.4$), and the results were therefore qualitatively summarized. The total weight of the studies with an upper CI limit < 0.25 was 61.4%, while the total weight of studies with an estimate > 0.25 was 15.8%. Consequently, this sensitivity analysis showed effects in the same direction and hence noninferiority at a margin of 0.25.

Metaanalyses without the Engleman 2002 study

To enable pooling of the primary studies' results, a metaanalysis without the Engleman study was performed as the 2nd type of sensitivity analysis. This is because the Engleman study's

results strongly deviated from the results of the other studies and were therefore largely responsible for heterogeneity, without any apparent medical or methodological reasons potentially explaining the deviating results of Engleman 2002. It was the only study with a statistically significant (and clinically relevant) effect in favour of MAD therapy.

The main analysis of the metaanalysis without the Engleman 2002 study was performed assuming a correlation of 0.6 for crossover studies which disregarded data dependency. The results were homogeneous, and pooling of the effect estimators was therefore possible. With an overall effect estimator of Hedges' g: 0.00; 95% CI: [−0.09; 0.09], noninferiority with a margin of 0.25 was found since the upper CI limit was below 0.25.

The sensitivity analysis assuming a correlation of 0.4 resulted in an overall effect estimate of Hedges' g: 0.00; 95% CI: [−0.09; 0.09], revealing noninferiority at a margin of 0.25.

The sensitivity analysis assuming a correlation of 0.8 resulted in an overall effect estimate of Hedges' g: 0.00; 95% CI: [−0.09; 0.09], resulting in noninferiority at a margin of 0.25.

Summary of results of the sensitivity analyses on noninferiority (ESS)

The main analysis, which assumed a correlation of 0.6, showed a result which was borderline for effects being in the same direction, depending on the rounding precision of Yamamoto 2019. The results of the associated sensitivity analyses assuming correlations of 0.4 and 0.8, respectively, showed MAD therapy to be noninferior to PAP therapy.

Under all 3 correlation assumptions (0.6, 0.4, and 0.8), both the shifting test (shifting to the disadvantage of MAD therapy) and the metaanalysis without Engleman 2002 (a study showing results to the disadvantage of MAD therapy) showed either effects in the same direction or, for pooled results, noninferiority at a margin of 0.25 (see Table 4). Hence, the results for noninferiority of MAD therapy in comparison with PAP therapy are considered robust.

Overall, there is consequently an indication of noninferiority of MAD therapy in comparison with PAP therapy.

Table 4: Results on the outcome of excessive daytime sleepiness – ESS (research question 2): Overview of sensitivity analyses on noninferiority

		Sensitivity analyses for the consideration of the statistical heterogeneity of study results		
		Main analysis	Sensitivity analysis 1: Shifting test	Sensitivity analysis 2: Omission of Engleman 2002
Sensitivity analyses for the consideration of disregarded data correlation in crossover studies	Main analysis: assumed correlation of 0.6	Effects marginally in the same direction (53.0% or 44.4%)[a]	0.07; 95% CI: [−0.05; 0.19][b]	0.00; 95% CI: [−0.09; 0.09][b]
	Sensitivity analysis 1: assumed correlation of 0.4	0.07; 95% CI: [−0.06; 0.20][b]	0.08; 95% CI: [−0.04; 0.20][b]	0.00; 95% CI: [−0.09; 0.09][b]
	Sensitivity analysis 2: assumed correlation of 0.8	Effects in the same direction (59.3%)[a]	Effects in the same direction (61.4%)[a]	0.00; 95% CI: [−0.09; 0.09][b]
a: Percentages indicate the total weight of the studies with an upper CI limit < 0.25. b: Hedges' g. CI: Confidence interval				

Subgroup analyses on excessive daytime sleepiness (ESS)

For excessive daytime sleepiness (ESS), the included studies tested for interaction regarding two characteristics of the device type (boil-and-bite versus custom-made appliance; monoblock versus twinblock design) as potential effect modifiers.

The test for interaction regarding the 1st device type characteristic (boil-and-bite versus custom-made appliance) showed no statistical significance (p = 0.111).

The test for interaction regarding the 2nd device type characteristic, appliance design (monoblock versus twinblock), failed to show any statistical significance as well (p = 0.243).

For research question 2, it was not possible to perform a subgroup analysis on excessive daytime sleepiness with regard to the adjustment of the degree of protrusion (customized degree of protrusion versus devices with pre-adjusted protrusion) because none of the studies used devices with pre-adjusted protrusion.

Further, for excessive daytime sleepiness (ESS), an interaction test was performed regarding the severity of OSA as a potential effect modifier. For the severity of OSA, however, the available evidence did not permit making the preplanned distinction between studies including persons with mild OSA versus those including persons with moderate OSA; a distinction was possible only between studies including persons with mild to moderate OSA versus studies including persons with mixed or unclear severity.

The test for interaction regarding the severity of OSA revealed no statistical significance (p = 0.503).

The included studies provided no usable results on other factors as potential effect modifiers. Therefore, it was not possible to perform any tests for interaction regarding other factors.

All things considered, for excessive daytime sleepiness, as assessed using ESS, there was no hint of effect modification by appliance type or severity of OSA.

Excessive daytime sleepiness – Berlin Questionnaire (BQ) for sleep apnoea: Single question on excessive daytime sleepiness

For excessive daytime sleepiness, 1 study (Ayra 2014) provided data of moderate qualitative certainty of results on the single question about excessive daytime sleepiness from the BQ. A statistically significant effect in favour of PAP therapy was found (MD: −0.5; 95% CI: [−0.82; −0.18]; p = 0.003).

Conclusion on benefit regarding excessive daytime sleepiness

The result of the single question regarding excessive daytime sleepiness from the BQ, which came from 1 study, was weighted lower than ESS in the benefit assessment. Although a statistically significant effect in favour of PAP therapy was observed for the single question, the metaanalysis' result on ESS showed no statistically significant effect. Since the ESS results stem from a metaanalysis of 13 studies, however, a superiority test regarding excessive daytime sleepiness revealed no hint of greater benefit or harm of MAD therapy in comparison with PAP therapy.

Conclusion on benefit regarding the noninferiority test on excessive daytime sleepiness

The overall consideration of the results of the main analysis, the shifting test, and the metaanalysis without Engleman 2002 as well as each of the sensitivity analyses on these metaanalyses with assumed correlations of 0.4 and 0.8, respectively, resulted in an indication of noninferiority of MAD therapy in comparison with PAP therapy when using a noninferiority margin of 0.25. The result of the single question on excessive daytime sleepiness from the BQ, which was based on 1 study, was weighted lower in the benefit assessment than the results from the ESS, which were based on a metaanalysis of 13 studies. The result of the single question therefore does not call into question the results of the noninferiority testing. Consequently, the overall assessment of excessive daytime sleepiness derived an indication of noninferiority of MAD therapy in comparison with PAP therapy.

4.3.4.3 Results on sleep quality

All instruments were weighted equally for deriving any benefit. Due to the differences in survey methods, no metaanalysis of results was compiled. Regarding sleep quality, 1 study (Arya 2014) provided usable data of moderate qualitative certainty of results for the comparison of MAD therapy versus PAP therapy on the basis of a single question from the BQ on fatigue upon awakening. No statistically significant effect was found (MD: −0.25; 95% CI: [−0.61; 0.11]; p = 0.169). Hence, for the BQ's single question on fatigue upon awakening, it was not possible to derive a hint of greater benefit or harm of MAD therapy compared to PAP therapy. In

addition, 1 study (El-Solh 2017) provided data of moderate qualitative certainty of results on PSQI Component 1 (subjective sleep quality) and on PSQI Component 5 (sleep disturbances). A statistically significant effect was not found for either of the two components (Component 1: MD: −0.13; 95% CI: [−0.61; 0.34]; p = 0.56; Component 5: MD: −0.13; 95% CI: [−0.51; 0.26]; p = 0.49). Hence, for the PSQI components, it was not possible to derive a hint of greater benefit or harm of MAD therapy compared to PAP therapy.

Conclusion on benefit regarding sleep quality

No statistically significant effect was found for any of the examined instruments or components. No further results were available. For sleep quality, it was consequently not possible to derive a hint of greater benefit or harm of MAD therapy compared to PAP therapy.

4.3.4.4 Results on cognitive performance (vigilance)

For the comparison of MAD therapy versus PAP therapy regarding cognitive performance (vigilance) as measured by errors of omission, 1 study (Barnes 2004) provided usable data of high qualitative certainty of results from the PVT, and 2 studies provided data of moderate qualitative certainty of results, 1 from the SteerClear test (Engleman 2002), and 1 from the AusEd Driving test (Phillips 2013).

Due to the absence of any statistically significant effect, it was not possible to draw any conclusions from the study of high qualitative certainty of results, Barnes 2004.

The studies of moderate qualitative certainty of results were then added, and the results from all 3 studies analysed together. Since the studies considerably differed in design and research question, particularly regarding the intervention, control, population, and instrument, an analysis was performed on the basis of a model with random effects as per Knapp and Hartung, initially including the data on the analysis type of “type 1 errors of omission” (called “lapses” in the study publication) for Phillips 2013.

The joint consideration of the 3 studies revealed a non-directed result since the pooled result was not statistically significant (Hedges’ g: 0.03; 95% CI: [−0.08; 0.13]; p = 0.404).

A qualitative evidence synthesis did not produce any different results because there were no effects in the same direction.

As an alternative analysis, for the Phillips 2013 study, the data on the analysis type of “type 2 errors of omission” (referred to as “crashes” in the study publication) were included in the metaanalysis.

Again, the joint consideration of the 3 studies revealed a non-directed result because the pooled result was not statistically significant (Hedges’ g: 0.01; 95% CI: [−0.22; 0.23]; p = 0.891).

A qualitative evidence synthesis did not produce any different results because there were no effects in the same direction.

Conclusion on benefit regarding the outcome of cognitive performance (vigilance)

All analysis types were weighted equally when deriving the benefit assessment. No statistically significant effect was found in either of the two metaanalyses on cognitive performance (vigilance). No further results were available. Consequently, in the overall analysis of cognitive performance (vigilance), it was not possible to derive a hint of greater benefit or harm of MAD therapy compared to PAP therapy.

4.3.4.5 Results on cognitive performance (executive functions)

For cognitive performance (executive functions), data on the Trail Making Test B were available from 3 studies (Barnes 2004, Engleman 2002, and Gagnadoux 2009), and on the PASAT, from 2 studies (Barnes 2004 and Engleman 2002). In addition, data on the F-A-S test / COWAT and digit span backwards were available from 1 study (Barnes 2004). Due to the difference in survey methods, no metaanalysis of the results of different instruments was compiled.

Results on the Trail Making Test B

In the comparison of MAD therapy versus PAP therapy, for cognitive performance (executive functions) as measured by the Trail Making Test B, 1 study (Barnes 2004) provided data of high qualitative certainty of results, and 2 studies (Engleman 2002, Gagnadoux 2009), data of moderate qualitative certainty of results.

Due to the absence of any statistically significant effect, it was not possible to draw any conclusions from the study of high qualitative certainty of results, Barnes 2004.

The studies of moderate qualitative certainty of results were then added, and the results from all 3 studies analysed together. Since the studies differed (considerably) in design and research question, particularly regarding the intervention, control, and population, an analysis was done based on a model with random effects as per Knapp and Hartung.

The joint consideration of the 3 studies revealed a non-directed result since the pooled result was not statistically significant (Hedges' g: 0.13; 95% CI: [−0.09; 0.35]; p = 0.132).

A qualitative evidence synthesis did not produce any different results because there were no effects in the same direction.

For the instrument "Trail Making Test B" measuring the outcome of cognitive performance (executive functions), it was therefore not possible to derive a hint of greater benefit or harm of MAD therapy compared to PAP therapy.

Results on the Paced Auditory Serial Addition Test (PASAT)

In the comparison of MAD therapy versus PAP therapy regarding cognitive performance (executive functions), the PASAT analysis was based on data of high qualitative certainty of results from 1 study (Barnes 2004) with 2 different method versions (interstimulus intervals of

1.2 and 2.4 seconds) and data of moderate qualitative certainty of results from 1 additional study (Engleman 2002), with an interstimulus interval of 2 seconds. The Barnes 2004 analysis with an interstimulus interval of 2.4 seconds was metaanalytically combined with the results of Engleman 2002.

Due to the absence of any statistically significant effect, it was not possible to draw any conclusions from the study of high qualitative certainty of results, Barnes 2004.

The study of moderate qualitative certainty of results, Engleman 2002, was then added, and the results from both studies were considered jointly. In the metaanalysis, an evaluation based on a model with fixed effect was conducted.

The combined analysis of both studies revealed a non-directed result because the pooled result was not statistically significant (Hedges' g: 0.03; 95% CI: [−0.11; 0.17]; p = 0.712).

The individual analysis of the Barnes 2004 results on PASAT with the interstimulus interval of 1.2 seconds showed a statistically significant but not clinically relevant effect in favour of MAD therapy (MD: −0.3; 95% CI: [−0.51; −0.09]; p < 0.004; Hedges' g: −0.38; 95% CI: [−0.65; −0.11]).

Regarding the outcome of cognitive performance (executive functions) as measured by the PASAT instrument, it was therefore not possible to derive a hint of greater benefit or harm of MAD therapy in comparison with PAP therapy.

Results on digit span backward

In the comparison of MAD therapy versus PAP therapy regarding cognitive performance (executive functions), 1 study (Barnes 2004) provided data of high qualitative certainty of results on digit span backward.

No statistically significant effect was found (MD: 0; 95% CI: [−0.28; 0.28]; p > 0.999).

Regarding the outcome of cognitive performance (executive functions) as measured by the digit span backward instrument, it was therefore not possible to derive a hint of greater benefit or harm of MAD therapy in comparison with PAP therapy.

Results on the F-A-S test / Controlled Oral Word Association Test (COWAT)

In the comparison of MAD therapy versus PAP therapy regarding cognitive performance (executive functions), 1 study (Barnes 2004) provided F-A-S test / COWAT data of high qualitative certainty of results.

No statistically significant effect was found (MD: −0.2; 95% CI: [−3.41; 3.01]; p = 0.902).

Regarding the outcome of cognitive performance (executive functions) as measured by the F-A-S test / COWAT instrument, it was therefore not possible to derive a hint of greater benefit or harm of MAD therapy in comparison with PAP therapy.

Conclusion on benefit regarding the outcome of cognitive performance (executive functions)

In the comparison of MAD therapy versus PAP therapy regarding cognitive performance (executive functions), results were available from the Trail Making Test B, PASAT, digit span backward, and F-A-S test / COWAT. All instruments were weighted equally for deriving any benefit. It was not possible to derive a hint of greater benefit or harm of MAD therapy compared to PAP therapy from any instrument. No further results were available. Overall, for cognitive performance (executive functions), it was therefore not possible to derive a hint of greater benefit or harm of MAD therapy in comparison with PAP therapy.

4.3.4.6 Results on depressive symptoms

For the comparison of MAD therapy versus PAP therapy regarding depressive symptoms, 1 study (Barnes 2004) provided data of moderate qualitative certainty of results from the BDI, and 2 studies (Engleman 2002, Hoekema 2006) provided data of moderate qualitative certainty of results from the Hospital Anxiety and Depression Scale (HADS).

A metaanalysis was conducted of the results from all 3 studies. Since the studies differed (considerably) in design and research question, particularly regarding the intervention, control, and population, an analysis was done based on a model with random effects as per Knapp and Hartung.

No statistically significant effect was found (Hedges' g: 0.08; 95% CI: [−0.39; 0.55]; p = 0.543).

A qualitative evidence synthesis did not produce any different results because there were no effects in the same direction.

Conclusion on benefit regarding the outcome of depressive symptoms

All instruments were weighted equally for deriving any benefit. A statistically significant effect was not found for BDI or HADS. No further results on this outcome were available. For depressive symptoms, it was therefore not possible to derive a hint of greater benefit or harm of MAD therapy compared to PAP therapy.

4.3.4.7 Results on anxiety symptoms

In the comparison of MAD therapy versus PAP therapy regarding anxiety symptoms, HADS data of moderate qualitative certainty of results were available from 2 studies (Engleman 2002, Hoekema 2006). In this case, an analysis was done based on a model with fixed effect.

No statistically significant effect was found (Hedges' g: 0.05; 95% CI: [−0.02; 0.12]; p = 0.151).

Conclusion on benefit regarding the outcome of anxiety symptoms

Regarding HADS, no statistically significant effect was found. No further results on this outcome were available. For anxiety symptoms, it was therefore not possible to derive a hint of greater benefit or harm of MAD therapy in comparison with PAP therapy.

4.3.4.8 Results on psychological symptoms

In the comparison of MAD therapy versus PAP therapy regarding psychological symptoms, 1 study (Aarab 2011) provided usable data of moderate qualitative certainty of results from SCL-90-R, and 1 study (Barnes 2004), from POMS. In each case, the total score was considered. Due to the difference in survey methods, no metaanalysis of the POMS and SCL-90-R instruments was performed.

No statistically significant effect was found for either the SCL-90-R total score (mean psychological burden) (MD: −4.8; 95% CI: [−39.91; 30.31]; p = 0.783) or for the POMS total score (MD: 3.4; 95% CI: [−1.92; 8.72]; p = 0.21).

Conclusion on benefit regarding the outcome of psychological symptoms

POMS and SCL-90-R were weighted equally in the benefit assessment, and only the total scores of the instruments were taken into account. No further results on this outcome were available. Since none of the instruments showed a statistically significant effect, it was not possible to derive, overall, a hint of greater benefit or harm of MAD therapy in comparison with PAP therapy regarding psychological symptoms.

4.3.4.9 Results on somatic symptoms

No results were available on somatic symptoms. Consequently, for this outcome, it was not possible to derive a hint of greater benefit or harm of MAD therapy in comparison with PAP therapy.

4.3.4.10 Results on cardiovascular morbidity

No results were available on cardiovascular morbidity. Consequently, for this outcome, it was not possible to derive a hint of greater benefit or harm of MAD therapy in comparison with PAP therapy.

4.3.4.11 Results on health-related quality of life

In the comparison of MAD therapy versus PAP therapy regarding health-related quality of life, data of moderate qualitative certainty of results were available from a total of 7 studies, of which 6 studies provided data from the SF-36 (Barnes 2004, El-Solh 2017, Engleman 2002, Hoekema 2006, Lam 2007, Phillips 2013), and 1 study, from the Nottingham Health Profile (NHP) (Gagnadoux 2009); the latter were combined with the SF-36 data in a metaanalysis.

For the SF-36 summary scores (SF-36 physical health summary score and SF-36 mental health summary score) and the SF-36 score of change in health, data were available from only 2 studies

each. Therefore, a metaanalysis of SF-36 and NHP was also conducted for the following subscores: SF-36 general health perceptions, SF-36 energy/fatigue + NHP loss of energy, SF-36 physical functioning + NHP limitatios in physical mobility, SF-36 bodily pain + NHP pain, SF-36 social functioning + NHP social isolation, SF-36 emotional well-being + NHP negative emotional reactions, SF-36 role limitations due to physical health problems, and SF-36 role limitations due to emotional health problems. In addition, an individual analysis was performed on the "sleep" subscore, on the basis of 1 study.

Since in all examined scores and subscores, the studies (substantially) differed in design and research question, particularly regarding the intervention, control, and population, a metaanalysis was conducted for each in the presence of 3 or 4 studies, on the basis of a model with random effects as per Knapp and Hartung.

In case of substantial heterogeneity, a qualitative analysis of results regarding the direction of the observed effects – but no metaanalysis – was conducted.

The individual analysis for the NHP subscore of sleep did not reveal any statistically significant effect (MD: −10.30; 95% CI: [−19.65; −0.95]; p = 0.031).

For the remaining scales and subscales, the metaanalysis showed no statistically significant and simultaneously clinically relevant effect.

A qualitative evidence synthesis in cases with 3 or 4 studies did not produce any different results because there were no effects in the same direction.

Conclusion on benefit regarding health-related quality of life

For SF-36, the summary scores as well as the subscores were used to derive any benefit, while for the NHP, all subscores were used. The metaanalyses on SF-36 and NHP as multidimensional instruments revealed no statistically significant and simultaneously clinically relevant effects; only the individual analysis of the NHP sleep subscore showed a statistically significant effect on the basis of 1 study. No further results on this outcome were available.

All things considered, for health-related quality of life, it was not possible to derive a hint of greater benefit or harm of MAD therapy in comparison with PAP therapy because the results of metaanalyses are associated with a greater certainty of results than the result of 1 subscore from 1 study.

4.3.4.12 Results on activities of daily living as well as participation in professional and social life

Regarding the comparison of MAD therapy versus PAP therapy, for activities of daily living and participation in professional and social life, FOSQ data of moderate qualitative certainty of results were available from 5 studies (Banhiran 2018, Barnes 2004, Engleman 2002, Hoekema 2006, Phillips 2013).

A metaanalysis of the results was then compiled on the basis of a model with random effects in accordance with the Knapp and Hartung method.

Due to heterogeneity ($p = 0.047$; $I^2 = 58.5\%$), meaningful pooling the results from the primary studies was not possible. A qualitative summary showed no effects in the same direction because none of the 5 studies provided a statistically significant and simultaneously clinically relevant result in favour of MAD therapy or PAP therapy (see General Methods 5.0 [116]); the total weight of the studies with clinically relevant effect was therefore 0%, and the prediction interval covered the zero effect.

Conclusion on benefit regarding activities of daily living as well as participation in professional and social life

From the FOSQ, only the total score was used to derive any benefit. No further results on this outcome were available. Since the qualitative summary regarding the FOSQ total score did not show any effects in the same direction, it was not possible to derive a hint of greater benefit or harm of MAD therapy in comparison with PAP therapy for activities of daily living or participation in professional and social life.

4.3.4.13 Results on serious adverse events and discontinuation due to adverse events

Generally, the studies surveyed or reported outcomes unsystematically. Specific information of moderate qualitative certainty of results on the occurrence of adverse events or discontinuation due to adverse events was found in a total of 7 studies (Aarab 2011, Banhiran 2018, El-Solh 2017, Ferguson 1996, Ferguson 1997, Lam 2007, and Tan 2002). Additional studies with reported results lacked data, for instance on the number of events, or it was unclear whether a reported discontinuation was due to adverse events or to other causes.

The reported specific adverse events of MAD therapy generally include, e.g. hypersalivation, temporomandibular joint pain, toothache, or occlusion impairment. Only in isolated cases did the description of events suggest that they were severe. In fact, studies repeatedly mentioned that the adverse events associated with MAD therapy were exclusively mild or only transient. The reported specific adverse events associated with PAP therapy generally included, e.g. dry mouth or throat, nose irritation or congestion, eye irritation or conjunctivitis, facial pain or facial lesions, and a feeling of suffocation. The available data from 1 study on serious adverse events did not suggest any higher frequency of serious adverse events in the MAD or PAP treatment group. No such events had occurred.

In terms of discontinuation due to adverse events, the 6 studies with usable data reported between 0 and 5 discontinuations due to adverse events in the intervention groups (MAD study arm), which corresponds to 0% to 11.8%, while in the control groups (PAP study arm), 0 to 6 discontinuations due to adverse events were reported for the control groups, corresponding to 0% to 28.6%. These results do not reveal any obviously higher frequency of discontinuations due to adverse events in either of the two therapy groups.

Apparently, no harm is discernible on the basis of the available data on the two outcomes of discontinuation due to adverse events and serious adverse events. On the basis of the available data, there is no hint of greater benefit or harm of MAD therapy in comparison with PAP therapy.

4.3.5 Results on AHI and ODI

Since AHI and ODI are neither patient-relevant outcomes nor validated surrogate outcomes, they were not used for the benefit assessment of MAD therapy. The results are discussed only as supplementary information in the details of the report.

4.4 Evidence map

Table 5 below shows the evidence map regarding patient-relevant outcomes.

Table 5: Evidence map regarding patient-relevant outcomes

		Research question 1	Research question 2
Mortality	**All-cause mortality / overall survival**	-	-
Morbidity	**Excessive daytime sleepiness**	⇑[a]	Superiority: ⇑⇓ Noninferiority: ⇑
	Sleep quality	⇔	⇔
	Cognitive performance (vigilance)	⇔	⇔
	Cognitive performance (executive functions)	⇔	⇔
	Depressive symptoms	⇔	⇔
	Anxiety symptoms	⇔	⇔
	Psychological symptoms	⇔	⇔
	Fatigue	⇗[b]	-
	Somatic symptoms (headache)	⇔	-
	Cardiovascular morbidity	-	-
	SAE	⇔	⇔
	Discontinuation due to AEs	⇔	⇔
Health-related quality of life	**Health-related quality of life**	⇔	⇔
	Activities of daily living as well as participation in professional and social life (FOSQ)	⇔	⇔

⇑: Indication of benefit of MAD therapy.
⇗: Hint of benefit of MAD therapy.
⇔: No hint, indication, or proof; homogeneous result.
⇑⇓: No hint, indication or proof; heterogeneous result.
-: No usable data reported.
a: Relates to the comparison of MAD therapy versus no treatment or placebo treatment not affecting the mandibular position.
b: Relates to the comparison of MAD vs. placebo device.
AE: adverse event; FOSQ: Functional Outcomes of Sleep Questionnaire; SAE: serious adverse event.

5 Classification of the assessment result

Many of the included studies used a crossover design, but several of them had disregarded the dependence structure of the data in their analyses. It cannot be determined to what extent and in which direction the results might consequently be biased. In the main analyses, the results for these studies were adjusted under the assumption of a correlation of 0.6, except for the supplementary discussions of AHI and ODI. Furthermore, for both research questions, sensitivity analyses assuming correlations of 0.4 and 0.8 were conducted for the outcome of excessive daytime sleepiness in order to cover a plausible value range concerning the size of the likely correlation.

Regarding research question 1, the identified studies stood out in terms of the wide range of different comparator interventions. Given that even of placebo devices and particularly of inactive devices were suspected to show effects, corresponding subgroup analyses were conducted for ESS. The analyses revealed an indication of benefit of MAD therapy with regard to the main symptom of OSA, excessive daytime sleepiness, only in comparison with no treatment or placebo treatment not affecting the mandibular position. This supports the hypothesis that at least some of the placebo devices and inactive devices were therapeutically effective, thus masking the effect of MAD therapy. Hence, the comparison with no treatment or placebo treatment not affecting the mandibular position was the only placebo condition in which therapeutic effect was to be suspected. For this reason, the result of the subgroup analysis by comparator intervention seemed more meaningful than that of the metaanalysis of all studies on excessive daytime sleepiness, with the comparison of MAD therapy versus no treatment or placebo treatment without impact on mandibular position being more relevant than the comparisons of MAD versus placebo device and MAD versus inactive device.

For research question 2, in addition to superiority testing regarding all patient-relevant outcomes, noninferiority testing was conducted regarding the main symptom of excessive daytime sleepiness. Superiority testing failed to reveal any hint of greater or lesser benefit of MAD therapy in comparison with PAP therapy for any of the examined patient-relevant outcomes. Hence, neither of the two therapies was found to be superior to the other with regard to these outcomes.

In noninferiority testing, the result of the main analysis for the main symptom of excessive daytime sleepiness was borderline for effects being in the same direction with the predefined noninferiority margin. Multiple sensitivity analyses were necessary to both adequately take into account the dependence structure of individual crossover studies and examine the heterogeneity caused particularly by one single study. Overall, the results were rated as robust.

MAD therapy is not inferior to PAP therapy with regard to the main symptom of OSA, excessive daytime sleepiness, and no relevant disadvantage was found with regard to other patient-relevant outcomes. This applies despite the fact that no (usable) data were available for overall mortality and cardiovascular morbidity. Nevertheless, it can be assumed that MAD

therapy is not relevantly inferior to PAP therapy with regard to these outcomes. No known evidence shows any advantage of PAP therapy over placebo or conservative treatment with regard to these outcomes in OSA patients. Instead, the available evidence actually suggests that PAP therapy does not statistically significantly reduce all-cause mortality or cardiovascular risk (see, e.g. [117-119]), and OSA therapy using MADs is unlikely to increase mortality or cardiovascular risk.

For both research questions, the evidence base did not permit any subgroup analyses of studies done exclusively on patients with mild OSA or exclusively on patients with moderate OSA. Therefore, no specific data were available to draw conclusions on any benefit of MAD therapy for patients with exclusively mild OSA or those with exclusively moderate OSA.

Some of the studies underlying the conclusions on benefit for either research question included patients with mild to moderate OSA, but others included only patients with severe OSA or mild to severe OSA (or unclear severity). However, there are no known reasons why the conclusion on benefit should not also apply to patients with mild to moderate OSA. Simultaneously, the data in no way suggested any lesser effects on daytime sleepiness for either research question in the two groups of interest versus in mixed populations. In a stratified analysis of the study results for patients with either mild to moderate OSA, or severe OSA only, or with mild to severe OSA (or unclear severity), forest plots showed no apparent differences (moreover, the corresponding test for interaction was not significant). The conclusions on benefit therefore also apply to individuals with mild to moderate OSA.

In this respect, it must be noted that MAD therapy is not suitable for all patients with mild or moderate OSA. For instance, MAD therapy is not suitable for edentulous patients or those with insufficient residual dentition, patients with extensive periodontitis or severe carious decay, or those with insufficient capability of mandibular protrusion or mouth opening [1,2,23].

Results on the outcomes of discontinuation due to adverse events and serious adverse events were reported only sporadically and unsystematically. Nevertheless, the available data on discontinuation due to adverse events and serious adverse events apparently reveal no discernible harm. In addition, it must be noted that no serious risks are discernible or expected on the basis of the mode of action of MADs and the fact that MAD therapy has been in use for many years. Hence, the available evidence on these outcomes does not contradict the conclusions on benefit of both research questions.

Complicating factors for both research questions were the large differences in the design of the examined MADs. In most studies, an individual device design was developed for MAD therapy as the experimental intervention. Particularly custom-made devices differed further in manufacturing methods and material. These major differences might have also contributed to the observed heterogeneity of results.

The vast majority of studies included for either research question (see Sections 4.2.1 and 4.3.1) examined custom-made, twinblock devices in which the set degree of protrusion was defined by individual adjustment. Few studies are currently available which involve a direct comparison of MADs with versus without adjustment [120].

A good evidence base is available only for excessive daytime sleepiness (and AHI). For the remaining outcomes, the evidence base is very patchy despite the large number of studies and large total number of examined patients. In this regard, it would be helpful if future studies relied on standardized tools to measure the individual outcomes. Future research should address which tools are well suited to this population.

In addition, no (usable) results were available at all for cardiovascular morbidity and mortality. This situation calls for studies with long follow-up periods, for which a parallel-group design is well suited, rather than the crossover design used by many of the included studies. Any future studies using a crossover design would benefit from a consideration of the dependence structure of the data, or at least a separate analysis by periods and testing for any carry-over effect.

For both research questions, ongoing studies exist which might be suitable for improving the evidence base in the foreseeable future.

The available information does not suggest publication bias for either of the two research questions.

6 Conclusion

Research question 1 – mandibular advancement device (MAD) therapy versus no treatment or placebo treatment

For excessive daytime sleepiness, an indication of benefit of MAD therapy in comparison with no treatment or with placebo treatment not affecting the mandibular position was found. For fatigue, there was a hint of benefit of MADs in comparison with placebo devices.

There was no hint of any benefit or harm of MAD therapy for the outcomes of sleep quality, cognitive performance (vigilance), cognitive performance (executive functions), depressive symptoms, anxiety symptoms, psychological symptoms, somatic symptoms (headache), health-related quality of life, activities of daily living as well as participation in professional and social life, serious adverse events, and discontinuation due to adverse events.

No data are available for the outcomes of overall mortality or overall survival and cardiovascular morbidity.

In summary, for research question 1, an advantage of MAD therapy regarding the main symptom of OSA, excessive daytime sleepiness, was found in comparison to either no treatment or placebo treatment not affecting the mandibular position. This advantage is not called into question by the results for other patient-relevant outcomes. For fatigue, an advantage of MAD therapy over a placebo device was found as well.

Research question 2 – MAD therapy versus positive airway pressure (PAP) therapy

For excessive daytime sleepiness, an indication of noninferiority of MAD therapy in comparison with PAP therapy was derived.

There was no hint of greater benefit or harm of MAD therapy for the outcomes of sleep quality, cognitive performance (vigilance), cognitive performance (executive functions), depressive symptoms, anxiety symptoms, psychological symptoms, health-related quality of life, activities of daily living as well as participation in professional and social life, serious adverse events, and discontinuation due to adverse events.

No (usable) data were available for the outcomes of all-cause mortality or overall survival, somatic symptoms, and cardiovascular morbidity.

In summary, for research question 2, MAD therapy was shown to be noninferior to PAP therapy regarding the main symptom of OSA, excessive daytime sleepiness, and simultaneously, no relevant disadvantage of MAD therapy versus PAP therapy was found for other patient-relevant outcomes.

References for English extract

Please see full final report for full reference list.

1. Deutsche Gesellschaft für Schlafforschung und Schlafmedizin. S3 Leitlinie Nicht erholsamer Schlaf/Schlafstörungen: Kapitel „Schlafbezogene Atmungsstörungen bei Erwachsenen“; Version 2.0 [online]. 08.2017 [Accessed: 17.10.2018]. URL: https://www.awmf.org/uploads/tx_szleitlinien/063-001l_S3_SBAS_2017-08_2.pdf.

2. Stuck BA, Maurer JT, Schlarb AA, Schredl M, Weeß HG. Praxis der Schlafmedizin: Diagnostik, Differenzialdiagnostik und Therapie bei Erwachsenen und Kindern. Berlin: Springer; 2018.

3. Lim J, Lasserson TJ, Fleetham J, Wright JJ. Oral appliances for obstructive sleep apnoea. Cochrane Database Syst Rev 2006; (1): CD004435.

4. Jordan AS, McSharry DG, Malhotra A. Adult obstructive sleep apnoea. Lancet 2014; 383(9918): 736-747.

5. Eckert DJ, Malhotra A. Pathophysiology of adult obstructive sleep apnea. Proc Am Thorac Soc 2008; 5(2): 144-153.

6. Malhotra RK, Kirsch DB, Kristo DA, Olson EJ, Aurora RN, Carden KA et al. Polysomnography for obstructive sleep apnea should include arousal-based scoring: an American Academy of Sleep Medicine position statement. J Clin Sleep Med 2018; 14(7): 1245-1247.

7. Epstein LJ, Kristo D, Strollo PJ Jr, Friedman N, Malhotra A, Patil SP et al. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. J Clin Sleep Med 2009; 5(3): 263-276.

8. Chervin RD. Sleepiness, fatigue, tiredness, and lack of energy in obstructive sleep apnea. Chest 2000; 118(2): 372-379.

9. George CFP. Sleep apnea, alertness, and motor vehicle crashes. Am J Respir Crit Care Med 2007; 176(10): 954-956.

10. Cheshire K, Engleman H, Deary I, Shapiro C, Douglas NJ. Factors impairing daytime performance in patients with sleep apnea/hypopnea syndrome. Arch Intern Med 1992; 152(3): 538-541.

11. Young T, Skatrud J, Peppard PE. Risk factors for obstructive sleep apnea in adults. JAMA 2004; 291(16): 2013-2016.

12. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM. Increased prevalence of sleep-disordered breathing in adults. Am J Epidemiol 2013; 177(9): 1006-1014.

13. Peppard PE, Young T, Palta M, Dempsey J, Skatrud J. Longitudinal study of moderate weight change and sleep-disordered breathing. JAMA 2000; 284(23): 3015-3021.

14. Lee RWW, Vasudavan S, Hui DS, Prvan T, Petocz P, Darendeliler MA et al. Differences in craniofacial structures and obesity in Caucasian and Chinese patients with obstructive sleep apnea. Sleep 2010; 33(8): 1075-1080.

15. Wetter DW, Young TB, Bidwell TR, Badr MS, Palta M. Smoking as a risk factor for sleep-disordered breathing. Arch Intern Med 1994; 154(19): 2219-2224.

16. Young T, Finn L, Austin D, Peterson A. Menopausal status and sleep-disordered breathing in the Wisconsin Sleep Cohort Study. Am J Respir Crit Care Med 2003; 167(9): 1181-1185.

17. Bixler EO, Vgontzas AN, Lin HM, Ten Have T, Rein J, Vela-Bueno A et al. Prevalence of sleep-disordered breathing in women: effects of gender. Am J Respir Crit Care Med 2001; 163(3 Pt 1): 608-613.

18. Patel SR, Larkin EK, Redline S. Shared genetic basis for obstructive sleep apnea and adiposity measures. Int J Obes (Lond) 2008; 32(5): 795-800.

19. Chi L, Comyn FL, Keenan BT, Cater J, Maislin G, Pack AI et al. Heritability of craniofacial structures in normal subjects and patients with sleep apnea. Sleep 2014; 37(10): 1689-1698.

20. Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. N Engl J Med 1993; 328(17): 1230-1235.

21. Heinzer R, Vat S, Marques-Vidal P, Marti-Soler H, Andries D, Tobback N et al. Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study. Lancet Respir Med 2015; 3(4): 310-318.

22. Gemeinsamer Bundesausschuss. Diagnostik und Therapie der schlafbezogenen Atmungsstörungen: zusammenfassender Bericht des Unterausschusses "Ärztliche Behandlung" des Gemeinsamen Bundesausschusses über die Beratungen von 1998 – 2004 zur Bewertung der Polygraphie und Polysomnographie im Rahmen der Differentialdiagnostik und Therapie der schlafbezogenen Atmungsstörungen gemäß §135 Abs.1 SGB V [online]. 27.01.2006 [Accessed: 31.08.2018]. URL: https://www.g-ba.de/downloads/40-268-234/2006-01-27-Abschluss-Polygraphie.pdf.

23. Deutsche Gesellschaft für Hals-Nasen-Ohren-Heilkunde, Kopf- und Hals-Chirurgie. Diagnostik und Therapie des Schnarchens des Erwachsenen: S3 Leitlinie (Langfassung) [online]. 03.2019 [Accessed: 16.01.2020]. URL: https://www.awmf.org/uploads/tx_szleitlinien/017-068l_S3_Diagnostik_Therapie_Schnarchen_Erwachsene_2019-8.pdf.

24. Deutsche Gesellschaft Zahnärztliche Schlafmedizin. Positionspapier der DGSZ zur Anwendung von Protrusionsschienen bei Erwachsenen mit schlafbezogenen Atmungsstörungen [online]. 01.09.2006 [Accessed: 13.09.2018]. URL: https://www.dgzs.de/_downloads/dgzs-positionspapier-protrusionsschienen-bei-sbas-2006.pdf.

25. Ferguson KA, Cartwright R, Rogers R, Schmidt-Nowara W. Oral appliances for snoring and obstructive sleep apnea: a review. Sleep 2006; 29(2): 244-262.

26. Dal-Fabbro C, Garbuio S, D'Almeida V, Cintra FD, Tufik S, Bittencourt L. Mandibular advancement device and CPAP upon cardiovascular parameters in OSA. Sleep and Breathing 2014; 18(4): 749-759.

27. Randerath WJ, Heise M, Hinz R, Ruehle KH. An individually adjustable oral appliance vs continuous positive airway pressure in mild-to-moderate obstructive sleep apnea syndrome. Chest 2002; 122(2): 569-575.

28. Banhiran W, Assanasen P, Nopmaneejumrudlers C, Nujchanart N, Srechareon W, Chongkolwatana C et al. Adjustable thermoplastic oral appliance versus positive airway pressure for obstructive sleep apnea. Laryngoscope 2018; 128(2): 516-522.

29. Lam B, Sam K, Mok WY, Cheung MT, Fong DY, Lam JC et al. Randomised study of three non-surgical treatments in mild to moderate obstructive sleep apnoea. Thorax 2007; 62(4): 354-359.

30. Hoekema A, Voors AA, Wijkstra PJ, Stegenga B, Van der Hoeven JH, Tol CG et al. Effects of oral appliances and CPAP on the left ventricle and natriuretic peptides. Int J Cardiol 2008; 128(2): 232-239.

31. Verse T, Dreher A, Heiser C, Herzog M, Maurer JT, Pirsig W et al. Leitlinie: „HNO-spezifische Therapie der obstruktiven Schlafapnoe bei Erwachsenen“ [online]. 09.2015 [Accessed: 17.10.2018]. URL: https://www.awmf.org/uploads/tx_szleitlinien/017-069l_S2e_Obstruktive_Schlafapnoe_Erwachsene_2015-12.pdf.

32. Sommer UJ, Heiser C, Gahleitner C, Herr RM, Hörmann K, Maurer JT et al. Tonsillectomy with uvulopalatopharyngoplasty in obstructive sleep apnea. Dtsch Arztebl Int 2016; 113(1-2): 1-8.

33. Zaghi S, Holty JE, Certal V, Abdullatif J, Guilleminault C, Powell NB et al. Maxillomandibular advancement for treatment of obstructive sleep apnea: a meta-analysis. JAMA Otolaryngol Head Neck Surg 2016; 142(1): 58-66.

34. Edmonds JC, Yang H, King TS, Sawyer DA, Rizzo A, Sawyer AM. Claustrophobic tendencies and continuous positive airway pressure therapy non-adherence in adults with obstructive sleep apnea. Heart Lung 2015; 44(2): 100-106.

35. Valentin A, Subramanian S, Quan SF, Berry RB, Parthasarathy S. Air leak is associated with poor adherence to autoPAP therapy. Sleep 2011; 34(6): 801-806.

36. Aarab G, Lobbezoo F, Hamburger HL, Naeije M. Oral appliance therapy versus nasal continuous positive airway pressure in obstructive sleep apnea: a randomized, placebo-controlled trial. Respiration 2011; 81(5): 411-419.

37. Barnes M, McEvoy RD, Banks S, Tarquinio N, Murray CG, Vowles N et al. Efficacy of positive airway pressure and oral appliance in mild to moderate obstructive sleep apnea. Am J Respir Crit Care Med 2004; 170(6): 656-664.

38. Aarab G, Lobbezoo F, Hamburger HL, Naeije M. Effects of an oral appliance with different mandibular protrusion positions at a constant vertical dimension on obstructive sleep apnea. Clin Oral Investig 2010; 14(3): 339-345.

39. Aarab G, Lobbezoo F, Heymans MW, Hamburger HL, Naeije M. Long-term follow-up of a randomized controlled trial of oral appliance therapy in obstructive sleep apnea. Respiration 2011; 82(2): 162-168.

40. Aarab G, Nikolopoulou M, Ahlberg J, Heymans MW, Hamburger HL, De Lange J et al. Oral appliance therapy versus nasal continuous positive airway pressure in obstructive sleep apnea: a randomized, placebo-controlled trial on psychological distress. Clin Oral Investig 2017; 21(7): 2371-2378.

41. Nikolopoulou M, Byraki A, Ahlberg J, Heymans MW, Hamburger HL, De Lange J et al. Oral appliance therapy versus nasal continuous positive airway pressure in obstructive sleep apnoea syndrome: a randomised, placebo-controlled trial on self-reported symptoms of common sleep disorders and sleep-related problems. J Oral Rehabil 2017; 44(6): 452-460.

42. VU University of Amsterdam. Oral appliance therapy in obstructive sleep apnea: study details [online]. In: ClinicalTrials.gov. 25.12.2014 [Accessed: 15.02.2019]. URL: https://clinicaltrials.gov/ct2/show/NCT00950495.

43. Andren A, Hedberg P, Walker-Engström ML, Wahlen P, Tegelberg A. Effects of treatment with oral appliance on 24-h blood pressure in patients with obstructive sleep apnea and hypertension: a randomized clinical trial. Sleep Breath 2013; 17(2): 705-712.

44. Blanco J, Zamarron C, Abeleira Pazos MT, Lamela C, Suarez Quintanilla D. Prospective evaluation of an oral appliance in the treatment of obstructive sleep apnea syndrome. Sleep Breath 2005; 9(1): 20-25.

45. Bloch KE, Iseli A, Zhang JN, Xie X, Kaplan V, Stoeckli PW et al. A randomized, controlled crossover trial of two oral appliances for sleep apnea treatment. Am J Respir Crit Care Med 2000; 162(1): 246-251.

46. Duran-Cantolla J, Crovetto-Martinez R, Alkhraisat MH, Crovetto M, Municio A, Kutz R et al. Efficacy of mandibular advancement device in the treatment of obstructive sleep apnea syndrome: a randomized controlled crossover clinical trial. Med Oral Patol Oral Cir Bucal 2015; 20(5): e605-e615.

47. Gagnadoux F, Pepin JL, Vielle B, Bironneau V, Chouet-Girard F, Launois S et al. Impact of mandibular advancement therapy on endothelial function in severe obstructive sleep apnea. Am J Respir Crit Care Med 2017; 195(9): 1244-1252.

48. Recoquillon S, Pepin JL, Vielle B, Andriantsitohaina R, Bironneau V, Chouet-Girard F et al. Effect of mandibular advancement therapy on inflammatory and metabolic biomarkers in patients with severe obstructive sleep apnoea: a randomised controlled trial. Thorax 26.10.2018 [Epub ahead of print].

49. University Hospital A. VASO-AM: impact of oral appliance therapy on endothelial function in obstructive sleep apnea (VASO-AM): study details [online]. In: ClinicalTrials.gov. 13.12.2016 [Accessed: 15.02.2019]. URL: https://clinicaltrials.gov/ct2/show/NCT01426607.

50. Godoy LBM, Palombini L, Poyares D, Dal-Fabbro C, Guimaraes TM, Klichouvicz PC et al. Long-term oral appliance therapy improves daytime function and mood in upper airway resistance syndrome patients. Sleep 2017; 40(12): 1-9.

51. Associação Fundo de Incentivo à Pesquisa. Efficacy of oral appliance for upper airway resistance syndrome: study details [online]. In: ClinicalTrials.gov. 15.07.2016 [Accessed: 12.06.2019]. URL: https://clinicaltrials.gov/ct2/show/NCT02636621.

52. Gotsopoulos H, Kelly JJ, Cistulli PA. Oral appliance therapy reduces blood pressure in obstructive sleep apnea: a randomized, controlled trial. Sleep 2004; 27(5): 934-941.

53. Gotsopoulos H, Chen C, Qian J, Cistulli PA. Oral appliance therapy improves symptoms in obstructive sleep apnea: a randomized, controlled trial. Am J Respir Crit Care Med 2002; 166(5): 743-748.

54. Naismith SL, Winter VR, Hickie IB, Cistulli PA. Effect of oral appliance therapy on neurobehavioral functioning in obstructive sleep apnea: a randomized controlled trial. J Clin Sleep Med 2005; 1(4): 374-380.

55. Hans MG, Nelson S, Luks VG, Lorkovich P, Baek SJ. Comparison of two dental devices for treatment of obstructive sleep apnea syndrome (OSAS). Am J Orthod Dentofacial Orthop 1997; 111(5): 562-570.

56. Johnston CD, Gleadhill IC, Cinnamond MJ, Gabbey J, Burden DJ. Mandibular advancement appliances and obstructive sleep apnoea: a randomized clinical trial. Eur J Orthod 2002; 24(3): 251-262.

57. Marklund M, Carlberg B, Forsgren L, Olsson T, Stenlund H, Franklin KA. Oral appliance therapy in patients with daytime sleepiness and snoring or mild to moderate sleep apnea: a randomized clinical trial. JAMA Intern Med 2015; 175(8): 1278-1285.

58. Rietz H, Franklin KA, Carlberg B, Sahlin C, Marklund M. Nocturnal blood pressure is reduced by a mandibular advancement device for sleep apnea in women: findings from secondary analyses of a randomized trial. J Am Heart Assoc 2018; 7(13): e008642.

59. Umeå University. Effects from a mandibular repositioning appliance in patients with obstructive sleep apnea and snoring: study details [online]. In: ClinicalTrials.gov. 18.01.2013 [Accessed: 15.02.2019]. URL: https://clinicaltrials.gov/ct2/show/NCT00477009.

60. Marklund M. Long-term efficacy of an oral appliance in early treated patients with obstructive sleep apnea. Sleep Breath 2016; 20(2): 689-694.

61. Mehta A, Qian J, Petocz P, Darendeliler MA, Cistulli PA. A randomized, controlled study of a mandibular advancement splint for obstructive sleep apnea. Am J Respir Crit Care Med 2001; 163(6): 1457-1461.

62. Petri N, Svanholt P, Solow B, Wildschiodtz G, Winkel P. Mandibular advancement appliance for obstructive sleep apnoea: results of a randomised placebo controlled trial using parallel group design. J Sleep Res 2008; 17(2): 221-229.

63. Nykøbing Falster County Hospital. Mandibular advancement device for obstructive sleep apnea: study details [online]. In: ClinicalTrials.gov. 15.09.2006 [Accessed: 15.02.2019]. URL: https://clinicaltrials.gov/ct2/show/NCT00243139.

64. Sjöholm TT, Polo OJ, Rauhala ER, Vuoriluoto J, Helenius HY. Mandibular advancement with dental appliances in obstructive sleep apnoea. J Oral Rehabil 1994; 21(5): 595-603.

65. De Britto Teixeira AO, Abi-Ramia LB, De Oliveira Almeida MA. Treatment of obstructive sleep apnea with oral appliances. Prog Orthod 2013; 14: 10.

66. Quinnell TG, Bennett M, Jordan J, Clutterbuck-James AL, Davies MG, Smith IE et al. A crossover randomised controlled trial of oral mandibular advancement devices for obstructive sleep apnoea-hypopnoea (TOMADO). Thorax 2014; 69(10): 938-945.

67. Sharples L, Glover M, Clutterbuck-James A, Bennett M, Jordan J, Chadwick R et al. Clinical effectiveness and cost-effectiveness results from the randomised controlled trial of oral mandibular advancement devices for obstructive sleep apnoea-hypopnoea (TOMADO) and long-term economic analysis of oral devices and continuous positive airway pressure. Health Technol Assess 2014; 18(67): 1-296.

68. Papworth Hospital NHS Foundation Trust. Oral mandibular advancement devices for obstructive sleep apnoea-hypopnoea [online]. In: ISRCTN Registry. 26.05.2016 [Accessed: 15.02.2019]. URL: http://www.isrctn.com/ISRCTN02309506.

69. Arya D, Singh SV, Tripathi A, Tripathi SK. A pilot study to compare patient perception of obstructive sleep apnea treatment with CPAP or appliance therapy. J Prosthet Dent 2014; 112(5): 1188-1193.

70. Mahidol University. Compare outcomes of SMG and CPAP in OSA: study details [online]. In: ClinicalTrials.gov. 09.05.2016 [Accessed: 15.02.2019]. URL: https://clinicaltrials.gov/ct2/show/NCT02766179.

71. Faculty of Medicine Siriraj Hospital. A randomized cross-over study of adjustable thermoplastic oral appliances and continuous positive airway pressure in treatment of patients with obstructive sleep apnea (RCT SMG and CPAP) [online]. In: Thai Clinical Trials Registry. 01.09.2014 [Accessed: 15.02.2019]. URL: http://www.clinicaltrials.in.th/index.php?tp=regtrials&menu=trialsearch&smenu=fulltext&task=search&task2=view1&id=1052.

72. Stanford University. Comparative outcomes management with electronic data technology (COMET) study (COMET): study details [online]. In: ClinicalTrials.gov. 17.05.2017 [Accessed: 06.11.2019]. URL: https://clinicaltrials.gov/ct2/show/NCT01461473.

73. Stanford University. Comparative outcomes management with electronic data technology (COMET) study (COMET): study results [online]. In: ClinicalTrials.gov. 17.05.2017 [Accessed: 15.02.2019]. URL: https://clinicaltrials.gov/ct2/show/results/NCT01461473.

74. De Vries GE, Hoekema A, Vermeulen KM, Claessen J, Jacobs W, Van der Maten J et al. Clinical- and cost-effectiveness of a mandibular advancement device versus continuous positive airway pressure in moderate obstructive sleep apnea. J Clin Sleep Med 2019; 15(10): 1477-1485.

75. University Medical Center Groningen. MRA therapy versus CPAP therapy in moderate OSAS (REST): study details [online]. In: ClinicalTrials.gov. 15.05.2018 [Accessed: 06.11.2019]. URL: https://clinicaltrials.gov/ct2/show/NCT01588275.

76. El-Solh AA, Homish GG, Ditursi G, Lazarus J, Rao N, Adamo D et al. A randomized crossover trial evaluating continuous positive airway pressure versus mandibular advancement device on health outcomes in veterans with posttraumatic stress disorder. J Clin Sleep Med 2017; 13(11): 1327-1335.

77. VA Office of Research and Development. A randomized cross over trial of two treatments for sleep apnea in veterans with post-traumatic stress disorder: study details [online]. In: ClinicalTrials.gov. 08.11.2017 [Accessed: 15.02.2019]. URL: https://clinicaltrials.gov/ct2/show/NCT01569022.

78. VA Western New York Healthcare System. A randomized cross over trial of two treatments for obstructive sleep apnea in veterans with post traumatic stress disorder: study details [online]. In: ClinicalTrials.gov. 17.02.2012 [Accessed: 15.02.2019]. URL: https://clinicaltrials.gov/ct2/show/NCT01535586.

79. VA Office of Research and Development. A randomized cross over trial of two treatments for sleep apnea in veterans with post-traumatic stress disorder: study results [online]. In: ClinicalTrials.gov. 08.11.2017 [Accessed: 15.02.2019]. URL: https://clinicaltrials.gov/ct2/show/results/NCT01569022.

80. Engleman HM, McDonald JP, Graham D, Lello GE, Kingshott RN, Coleman EL et al. Randomized crossover trial of two treatments for sleep apnea/hypopnea syndrome: continuous positive airway pressure and mandibular repositioning splint. Am J Respir Crit Care Med 2002; 166(6): 855-859.

81. Ferguson KA, Ono T, Lowe AA, Keenan SP, Fleetham JA. 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(5): 1269-1275.

82. Ferguson KA, Ono T, Lowe AA, Al-Majed S, Love LL, Fleetham JA. A short-term controlled trial of an adjustable oral appliance for the treatment of mild to moderate obstructive sleep apnoea. Thorax 1997; 52(4): 362-368.

83. Gagnadoux F, Fleury B, Vielle B, Petelle B, Meslier N, N'Guyen XL et al. Titrated mandibular advancement versus positive airway pressure for sleep apnoea. Eur Respir J 2009; 34(4): 914-920.

84. Trzepizur W, Gagnadoux F, Abraham P, Rousseau P, Meslier N, Saumet JL et al. Microvascular endothelial function in obstructive sleep apnea: impact of continuous positive airway pressure and mandibular advancement. Sleep Med 2009; 10(7): 746-752.

85. University Hospital Angers. Treatment of sleep apnea syndrome with nCPAP versus oral appliance: study details [online]. In: ClinicalTrials.gov. 28.10.2015 [Accessed: 12.06.2019]. URL: https://clinicaltrials.gov/ct2/show/NCT00152672.

86. Glos M, Penzel T, Schoebel C, Nitzsche GR, Zimmermann S, Rudolph C et al. Comparison of effects of OSA treatment by MAD and by CPAP on cardiac autonomic function during daytime. Sleep Breath 2016; 20(2): 635-646.

87. Charité - Universitätsmedizin Berlin. Effectiveness of OSA treatment by MAD and CPAP on cardiac autonomic function at daytime [online]. In: Deutsches Register Klinischer Studien. [Accessed: 15.02.2019]. URL: http://www.drks.de/DRKS00007772.

88. Hoekema A. Efficacy and comorbidity of oral appliances in the treatment of obstructive sleep apnea-hypopnea: a systematic review and preliminary results of a randomized trial. Sleep Breath 2006; 10(2): 102-103.

89. Hoekema A, Stegenga B, Bakker M, Brouwer WH, De Bont LG, Wijkstra PJ et al. Simulated driving in obstructive sleep apnoea-hypopnoea: effects of oral appliances and continuous positive airway pressure. Sleep Breath 2007; 11(3): 129-138.

90. Hoekema A, Stel AL, Stegenga B, Van der Hoeven JH, Wijkstra PJ, Van Driel MF et al. Sexual function and obstructive sleep apnea-hypopnea: a randomized clinical trial evaluating the effects of oral-appliance and continuous positive airway pressure therapy. J Sex Med 2007; 4(4): 1153-1162.

91. Hoekema A, Stegenga B, Wijkstra PJ, Van der Hoeven JH, Meinesz AF, De Bont LG. Obstructive sleep apnea therapy. J Dent Res 2008; 87(9): 882-887.

92. Doff MH, Hoekema A, Pruim GJ, Huddleston Slater JJ, Stegenga B. Long-term oral-appliance therapy in obstructive sleep apnea: a cephalometric study of craniofacial changes. J Dent 2010; 38(12): 1010-1018.

93. Doff MH, Veldhuis SK, Hoekema A, Slater JJ, Wijkstra PJ, De Bont LG et al. Long-term oral appliance therapy in obstructive sleep apnea syndrome: a controlled study on temporomandibular side effects. Clin Oral Investig 2012; 16(3): 689-697.

94. Doff MH, Hoekema A, Wijkstra PJ, Van der Hoeven JH, Huddleston Slater JJ, De Bont LG et al. Oral appliance versus continuous positive airway pressure in obstructive sleep apnea syndrome: a 2-year follow-up. Sleep 2013; 36(9): 1289-1296.

95. Doff MH, Finnema KJ, Hoekema A, Wijkstra PJ, De Bont LG, Stegenga B. Long-term oral appliance therapy in obstructive sleep apnea syndrome: a controlled study on dental side effects. Clin Oral Investig 2013; 17(2): 475-482.

96. University Medical Center Groningen. Management of the obstructive sleep apnea-hypopnea syndrome: oral appliance versus continuous positive airway pressure therapy [online]. In: Netherlands Trial Register. [Accessed: 15.02.2019]. URL: https://www.trialregister.nl/trial/75.

97. Phillips CL, Grunstein RR, Darendeliler MA, Mihailidou AS, Srinivasan VK, Yee BJ et al. Health outcomes of continuous positive airway pressure versus oral appliance treatment for obstructive sleep apnea: a randomized controlled trial. Am J Respir Crit Care Med 2013; 187(8): 879-887.

98. Sutherland K, Phillips CL, Davies A, Srinivasan VK, Dalci O, Yee BJ et al. CPAP pressure for prediction of oral appliance treatment response in obstructive sleep apnea. J Clin Sleep Med 2014; 10(9): 943-949.

99. Woolcock Institute of Medical Research. Randomised controlled trial of the effect of mandibular advancement splint (MAS) versus positive airway pressure (PAP) therapy on blood pressure in obstructive sleep apnoea [online]. In: Australian New Zealand Clinical Trials Registry. 09.11.2015 [Accessed: 15.02.2019]. URL: https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?ACTRN=12607000289415.

100. Schütz TC, Cunha TC, Moura-Guimaraes T, Luz GP, Ackel-D'Elia C, Alves Eda S et al. Comparison of the effects of continuous positive airway pressure, oral appliance and exercise training in obstructive sleep apnea syndrome. Clinics 2013; 68(8): 1168-1174.

101. Associação Fundo de Incentivo à Pesquisa. Continuous positive airway pressure (CPAP), oral appliance and physical exercise in the obstructive sleep apnea: study results [online]. In: ClinicalTrials.gov. 03.05.2013 [Accessed: 15.02.2019]. URL: https://clinicaltrials.gov/ct2/show/results/NCT01289392.

102. Tan YK, L'Estrange PR, Luo YM, Smith C, Grant HR, Simonds AK et al. Mandibular advancement splints and continuous positive airway pressure in patients with obstructive sleep apnoea: a randomized cross-over trial. Eur J Orthod 2002; 24(3): 239-249.

103. Yamamoto U, Nishizaka M, Tsuda H, Tsutsui H, Ando SI. Crossover comparison between CPAP and mandibular advancement device with adherence monitor about the effects on endothelial function, blood pressure and symptoms in patients with obstructive sleep apnea. Heart Vessels 2019; 34(10): 1692-1702.

104. Sleep Apnea Center, Kyushu University Hospital. Comparison of endothelial function with the mandibular advancement device (MAD) versus continuous positive airway pressure (CPAP) among patients with moderate sleep apnea syndrome: multicenter, open label, prospective randomized non-inferiority crossover trial [online]. In: UMIN Clinical Trials Registry. 03.08.2018 [Accessed: 08.01.2020]. URL: https://upload.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000017028.

105. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. General methods: version 5.0 [online]. 10.07.2017 [Zugriff: 06.06.2018]. URL: https://www.iqwig.de/download/General-Methods_Version-5-0.pdf.

106. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ et al. Cochrane handbook for systematic reviews of interventions: version 6.0 [online]. 07.2019 [Accessed: 03.01.2020]. URL: https://training.cochrane.org/handbook.

107. Vroegop AV, Vanderveken OM, Van de Heyning PH, Braem MJ. Effects of vertical opening on pharyngeal dimensions in patients with obstructive sleep apnoea. Sleep Med 2012; 13(3): 314-316.

108. Antic NA, Buchan C, Esterman A, Hensley M, Naughton MT, Rowland S et al. A randomized controlled trial of nurse-led care for symptomatic moderate-severe obstructive sleep apnea. Am J Respir Crit Care Med 2009; 179(6): 501-508.

109. Bloch KE, Huber F, Furian M, Latshang TD, Lo Cascio CM, Nussbaumer-Ochsner Y et al. Autoadjusted versus fixed CPAP for obstructive sleep apnoea: a multicentre, randomised equivalence trial. Thorax 2018; 73(2): 174-184.

110. Chai-Coetzer CL, Antic NA, Rowland LS, Reed RL, Esterman A, Catcheside PG et al. Primary care vs specialist sleep center management of obstructive sleep apnea and daytime sleepiness and quality of life: a randomized trial. JAMA 2013; 309(10): 997-1004.

111. Chai-Coetzer CL, Antic NA, Hamilton GS, McArdle N, Wong K, Yee BJ et al. Physician decision making and clinical outcomes with laboratory polysomnography or limited-channel sleep studies for obstructive sleep apnea: a randomized trial. Ann Intern Med 2017; 166(5): 332-340.

112. Corral J, Sanchez-Quiroga MA, Carmona-Bernal C, Sanchez-Armengol A, De la Torre AS, Duran-Cantolla J et al. Conventional polysomnography is not necessary for the management of most patients with suspected obstructive sleep apnea: noninferiority, randomized controlled trial. Am J Respir Crit Care Med 2017; 196(9): 1181-1190.

113. Sanchez-Quiroga MA, Corral J, Gomez-de-Terreros FJ, Carmona-Bernal C, Asensio-Cruz MI, Cabello M et al. Primary care physicians can comprehensively manage sleep apnea patients: a non-inferiority randomized controlled trial. Am J Respir Crit Care Med 01.09.2018 [Epub ahead of print].

114. Tarraubella N, De Batlle J, Nadal N, Castro-Grattoni AL, Gomez S, Sanchez-de-la-Torre M et al. GESAP trial rationale and methodology: management of patients with suspected obstructive sleep apnea in primary care units compared to sleep units. NPJ Prim Care Respir Med 2017; 27(1): 8.

115. Tarraubella N, Sanchez-de-la-Torre M, Nadal N, De Batlle J, Benitez I, Cortijo A et al. Management of obstructive sleep apnoea in a primary care vs sleep unit setting: a randomised controlled trial. Thorax 2018; 73(12): 1152-1160.

116. Schwartz M, Acosta L, Hung YL, Padilla M, Enciso R. Effects of CPAP and mandibular advancement device treatment in obstructive sleep apnea patients: a systematic review and meta-analysis. Sleep Breath 2018; 22(3): 555-568.

117. Elbourne DR, Altman DG, Higgins JPT, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. Int J Epidemiol 2002; 31(1): 140-149.

118. Sivaramakrishnan G, Sridharan K. A systematic review on the effectiveness of titratable over nontitratable mandibular advancement appliances for sleep apnea. J Indian Prosthodont Soc 2017; 17(4): 319-324.

119. Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455.

120. Lefebvre C, Manheimer E, Glanville J. Searching for studies [online]. In: Higgings JPT, Green S (Ed). Cochrane handbook for systematic reviews of interventions: version 5.1.0. 03.2011 [Accessed: 05.09.2018]. URL: http://handbook-5-1.cochrane.org/chapter_6/6_searching_for_studies.htm.

The full report (German version) is published under https://www.iqwig.de/en/projects-results/projects/non-drug-interventions/n-projekte/n18-03-mandibular-advancement-devices-in-mild-to-moderate-obstructive-sleep-apnoea-in-adults.9673.html.

Appendix A – Search strategies

A.1 – Searches in bibliographic databases

1. MEDLINE

Search interface: Ovid

- Ovid MEDLINE(R) 1946 to August Week 2 2019,
- Ovid MEDLINE(R) Daily Update August 20, 2019,
- Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations 1946 to August 20, 2019,
- Ovid MEDLINE(R) Epub Ahead of Print August 20, 2019

The following filters were adopted:

- Systematic review: Wong [119] – High specificity strategy
- RCT: Lefebvre [120] – Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision)

#	**Searches**
1	exp Sleep Apnea, Obstructive/ or *Sleep Apnea Syndromes/
2	(obstructive* adj1 sleep* adj1 apn?ea*).ti,ab.
3	(sleep* adj1 apn?ea* adj3 syndrome*).ti,ab.
4	or/1-3
5	exp Orthodontic Appliances/ or Mandibular Advancement/
6	((oral* or dental* or mandibular*) adj3 (appliance* or advancement* or device* or splint*)).ti,ab.
7	5 or 6
8	randomized controlled trial.pt.
9	controlled clinical trial.pt.
10	(randomized or placebo or randomly or trial or groups).ab.
11	drug therapy.fs.
12	or/8-11
13	exp animals/ not humans.sh.
14	12 not 13
15	Cochrane database of systematic reviews.jn.
16	meta analysis.pt.
17	(search or MEDLINE or systematic review).tw.
18	or/15-17
19	14 or 18
20	and/4,7,19
21	20 not (comment or editorial).pt.

2. PubMed

Search interface: NLM

- PubMed – as supplied by publisher
- PubMed – in process
- PubMed – pubmednotmedline

Search	Query
#1	Search (obstructive* [TIAB] AND sleep* [TIAB] AND (apnea* [TIAB] OR apnoea* [TIAB]))
#2	Search (sleep* [TIAB] AND (apnea* [TIAB] OR apnoea* [TIAB]) AND syndrome* [TIAB])
#3	Search (#1 OR #2)
#4	Search ((oral* [TIAB] OR dental* [TIAB] OR mandibular* [TIAB]) AND (appliance* [TIAB] OR advancement* [TIAB] OR device* [TIAB] OR splint* [TIAB]))
#5	Search (clinical trial*[TIAB] OR random*[TIAB] OR placebo[TIAB] OR trial[TI])
#6	Search (search[TIAB] OR meta analysis[TIAB] OR MEDLINE[TIAB] OR systematic review[TIAB])
#7	Search (#5 OR #6)
#8	Search (#3 AND #4 AND #7)
#9	Search (#8 NOT Medline [SB])

3. Embase

Search interface: Ovid

- Embase 1974 to 2019 August 20

The following filters were adopted:

- Systematic review: Wong [119] – High specificity strategy;
- RCT: Wong [119] – Strategy minimizing difference between sensitivity and specificity

#	Searches
1	sleep disordered breathing/ or sleep apnea syndrome/ or sleep apnea, obstructive/
2	(obstructive* adj1 sleep* adj1 apn?ea*).ti,ab.
3	(sleep* adj1 apn?ea* adj3 syndrome*).ti,ab.
4	or/1-3
5	mandible reconstruction/ or orthodontic device/ or mandibular advancement/ or mandibular advancement device/

6	((oral* or dental* or mandibular*) adj3 (appliance* or advancement* or device* or splint*)).ti,ab.
7	or/5-6
8	(random* or double-blind*).tw.
9	placebo*.mp.
10	or/8-9
11	(meta analysis or systematic review or MEDLINE).tw.
12	or/10-11
13	and/4,7,12
14	13 not medline.cr.
15	14 not (exp animal/ not exp human/)
16	15 not (Conference Abstract or Conference Review or Editorial).pt.

4. The Cochrane Library

Search interface: Wiley

- Cochrane Database of Systematic Reviews: Issue 8 of 12, August 2019
- Cochrane Central Register of Controlled Trials: Issue 8 of 12, August 2019

ID	**Search**
#1	[mh "Sleep Apnea, Obstructive"]
#2	[mh ^"Sleep Apnea Syndromes"]
#3	(obstructive* near/1 sleep* near/1 apn*ea*):ti,ab
#4	(sleep* near/1 apn*ea* near/3 syndrome*):ti,ab
#5	#1 or #2 or #3 or #4
#6	[mh "Orthodontic Appliances"]
#7	[mh ^"Mandibular Advancement"]
#8	((oral* or dental* or mandibular*) near/3 (appliance* or advancement* or device* or splint*)):ti,ab
#9	#6 or #7 or #8
#10	#5 and #9 in Cochrane Reviews, Cochrane Protocols
#11	#5 and #9 in Trials

5. Health Technology Assessment Database

Search interface: Centre for Reviews and Dissemination

Line	Search
1	MeSH DESCRIPTOR Sleep Apnea, Obstructive EXPLODE ALL TREES
2	MeSH DESCRIPTOR Sleep Apnea Syndromes
3	(obstructive* AND sleep* AND (apnea* OR apnoea*))
4	(sleep* AND (apnea* OR apnoea*) AND syndrome*)
5	#1 OR #2 OR #3 OR #4
6	MeSH DESCRIPTOR Orthodontic Appliances EXPLODE ALL TREES
7	MeSH DESCRIPTOR Mandibular Advancement
8	((oral* OR dental* OR mandibular*) AND (appliance* AND advancement* OR device* OR splint*))
9	#6 OR #7 OR #8
10	#5 AND #9
11	(#10) IN HTA

A.2 – Searches in study registries

1. ClinicalTrials.gov

Provider: U.S. National Institutes of Health

- URL: http://www.clinicaltrials.gov
- Type of search: Advanced Search

Search strategy
(sleep apnea syndrome OR obstructive sleep apnea OR upper airway resistance syndrome) AND (oral OR dental OR mandibular) AND (appliance OR advancement OR device OR splint)

2. International Clinical Trials Registry Platform Search Portal

Provider: World Health Organization

- URL: http://apps.who.int/trialsearch/
- Type of search: Standard Search

Search strategy
oral appliance* OR oral advancement* OR oral device* OR oral splint* OR dental appliance* OR dental advancement* OR dental device* OR dental splint* OR mandibular appliance* OR mandibular advancement* OR mandibular device* OR mandibular splint*