Quality Assessment of Systematic Reviews on the Efficacy of Oral Appliance Therapy for Adult and Pediatric Sleep-Disordered Breathing

Thikriat S. Al-Jewair, BDS, MBA, MSc, MS, FRDCDC; Balgis O. Gaffar, BDS, MS; Carlos Flores-Mir, DDS, DSc, FRDCDC

1Clinical Assistant Professor and Program Director, Department of Orthodontics, State University of New York at Buffalo, Buffalo, NY; 2Lecturer, Department of Preventive Dental Sciences, College of Dentistry, University of Dammam, Saudi Arabia; 3Professor, Orthodontic Graduate Program Director, School of Dentistry, Faculty of Medicine and Dentistry, University of Alberta, Edmonton Clinic Health Academy, University of Alberta, Edmonton, Alberta Canada

INTRODUCTION

Loud snoring and obstructive sleep apnea (OSA) are common sleep-disordered breathing (SDB) conditions that can affect growing and nongrowing individuals. OSA is characterized by repetitive partial (hypopnea) or complete (apnea) upper airway obstructions lasting at least 10 sec and by snoring during sleep. The prevalence in adults ranges from 9% to 24% for OSA and from 2% to 4% for OSA when daytime symptoms are considered. OSA has been linked to cardiovascular disease and hypertension in adults. Disorders, childhood obesity as a result of the deregulation of appetite-modulating neuropeptides, and delayed development are outcomes of SDB in children and adolescents.

Multiple factors have been identified in the etiology of SDB/OSA. These included structural constrictions of the soft tissue space surrounding the oropharynx and nasopharynx and its lumen (i.e., hypertrophied tonsils and/or adenoids), constrictions surrounding the nasal cavity, neuromuscular abnormalities of the upper airway, hyperdivergent craniofacial growth patterns, genetic markers (i.e., tumor necrosis factor-alpha), early life programming (i.e., exposure to secondhand cigarette smoking and poor nutrition), and prenatal factors such as hormonal and physiological changes during gestation, in addition to other factors.

Pediatric and adult OSA are serious disorders. Untreated OSA has been linked to cardiovascular disease and hypertension in adults. As well, attention deficit hyperactivity disorders, childhood obesity as a result of the deregulation of appetite-modulating neuropeptides, and delayed development are outcomes of SDB in children and adolescents.

To date, the first line of treatment for children with existing disease and/or at high risk for OSA is tonsillectomy and/or adenoidectomy. Other treatments include weight loss, intranasal corticosteroids, and continuous positive airway pressure, which has been found to limit craniofacial development. Oral appliances (OAs) such as maxillary expanders and mandibular orthopedic appliances have also been suggested to improve SDB outcomes.

Treatments for adult SDB include both surgical and nonsurgical interventions. Non-surgical interventions include one or...
more of the following: adopting a healthier lifestyle (i.e., losing weight, avoiding alcohol, and positional therapy), CPAP therapy, or OA therapy. OA therapy has shown improved effectiveness along with a higher level of compliance compared to the gold standard CPAP therapy. The American Academy of Sleep Medicine has set OAs treatment management parameters.

When properly indicated, OAs act by repositioning the mandible with the supratohyoid and genioglossal muscles attachments and tongue forward and inferiorly. This anterior and inferior movement of the mandible reduces the gravitational effect of the tongue, and stretches the palatoglossal and palatopharyngeal arches. This increases the dimensions of the velopharyngeal airway, thus potentially improving upper airway patency and preventing obstruction. There are different OA design modifications that may include tongue retaining, soft palate lifting, and acrylic-based mandibular advancing devices.

Reliable scientific evidence on safety and efficacy comes from systematic reviews (SRs) of high-quality randomized controlled trials (RCTs). Evidence from these reviews supports stronger evidence-based guidelines for best clinical practices. With the growing number of published SRs in the literature, careful assessment of their methodological quality becomes important since evidence from low methodological quality SRs with high risk of bias is questionable and in some cases invalid.

Many SRs have been published on the efficacy of OA therapy in adult SDB, but only few about pediatric SDB. Yet, variability exists in the methodological quality of these reviews, which affects the applicability of their results and the subsequent interpretation of these results by professionals. Many factors have been identified that contribute to this variability, including types of included studies, reporting transparency, and methods employed to combine available data.

For a life-threatening condition, such as severe SDB forms (namely OSA), extra efforts should be undertaken to identify sound high quality evidence-based guidelines to guide best practice for the management of SDB. In an effort to assess the quality of published SRs on the efficacy of OAs in adult severe SDB (OSA), Johal et al. evaluated eight reviews/clinical practice guidelines comparing OAs with different modalities including no treatment. They reported variable quality with only two of the SRs being of high quality. More studies were published since their last search in 2013 and no study has evaluated the quality of SRs about pediatric OSA. Therefore, the objective of this study was to assess the methodological quality of SRs on the efficacy of OA therapy in the management of pediatric and adult SDB.

Methods

The preferred reported items for systematic reviews and meta-analyses (PRISMA) criteria were followed during the reporting of this methodological quality assessment study.

Study Inclusion/Exclusion

Studies considered for inclusion were SRs with or without meta-analyses (MAs) published in any language. Only reviews that used the term “systematic review” or “meta-analysis” in their titles or their methodologies clearly indicated that a systematic review had been undertaken were included. Narrative reviews in which only systematic literature searching methods were used were excluded. The reviews had to evaluate the efficacy or effectiveness of OA therapy in the treatment of SDB including OSA of all levels of severity in children and/or adults. All types of OAs were considered without restrictions. Reviews comparing OAs with other interventions were also considered for inclusion. Primary studies, descriptive reviews, case reports, case series, editorials, commentaries, Task Force reports, clinical guidelines, and practice parameters were all excluded.

Data Sources

Multiple electronic databases (PubMed, Ovid MEDLINE, EMBASE, Evidence Based Medicine Reviews, Cumulative Index to Nursing & Allied Health (CINAHL), and ProQuest Digital Dissertation) were searched for articles published from the database inception until January 2016. In addition, the Health Technology Assessment database, the American Dental Association Center for Evidence-Based Dentistry (ADA-EBD) website, and the World Wide Web using Google Scholar were searched. Reference lists of retrieved articles were also crosschecked for potential articles that may have been missed. A combination of Keywords and medical subject headings (MeSH) were used based on the database (supplemental material).

Two independent reviewers (T.J. and B.G.) selected the studies independently. All titles and abstracts were reviewed to identify potential studies. Articles that fulfilled the inclusion criteria or had incomplete information at this stage underwent full-text review to carefully assess their eligibility. Data were then extracted using a customized data abstraction sheet. Any disagreement was resolved through discussion.

Methodological Quality Assessment

The same two reviewers independently assessed the methodological quality of the studies using the Assessment of Multiple Systematic Reviews (AMSTAR) measurement tool. The validity and reliability of the AMSTAR has been established in previous studies.

The AMSTAR tool assesses 11 domains in systematic reviews of intervention: a priori design, study selection and data extraction in duplicate, comprehensive literature search (at least two electronic database literature sources and one supplementary source), search for reports regardless of their publication status (i.e., gray literature), list of included/excluded studies, presentation of the characteristics of the included studies, documentation of methodological quality assessment, well-supported conclusions that are based on the quality of evidence included, appropriate methods to combine individual findings of the selected studies or if discussed reasons for not combining studies, assessment of the likelihood of publication bias, and documentation of conflicts of interest and sources of funding in both the systematic review and the included studies. The possible responses for each item are “yes”, “no”, “cannot answer”, or “not applicable”. One point was given for every “yes” response and no points if the other responses were selected.
Total scores were then calculated with the maximum possible score being eleven. A systematic review was described as high quality if scored (9–11), medium or low quality if scored (5–8) and (< 4) respectively. Disagreement in scoring between investigators was discussed openly and when required, a third investigator was consulted (C.F.M).

**RESULTS**

A total of 387 studies were identified and assessed for inclusion. After exclusion on the title and abstract stages, 44 articles were retrieved for full review. Two further articles were identified through reference search of retrieved articles. Twenty-nine articles\(^{13,21-48}\) were later excluded after full-text review for different reasons (Table 1). Therefore, only 17 articles met the inclusion criteria set for this study (Figure 1). Thirteen reviews were on adult OSA (2 SRs and 11 SRs with MAs),\(^ {1,49-60}\) and four reviews were on pediatric OSA (3 SRs and 1 SR with MA).\(^ {12,15,61,62}\) Key methodological and descriptive characteristics of the included reviews are presented in Tables 2 and 3.

All the included reviews were published between 2004 and 2016 and were in English language except for one review\(^ {57}\) in Chinese. The number of included studies in each review ranged from 1 to 71 studies with different study designs. As expected, RCTs was the most commonly sought and included study type.

The studies included in the adult OSA reviews had the OA groups compared to one or more of the following groups: placebo, CPAP, surgery, another OA of different design, or pretreatment and posttreatment with only one OA type. With respect to the pediatric OSA reviews, the comparison groups included no treatment, placebo, and pretreatment and posttreatment with only one OA type. Two reviews\(^ {12,15}\) included studies on maxillary expansion with and without concurrent use of fixed orthodontic appliances. Another review\(^ {15}\) included five RCTs that assessed seven independent interventions with only one of those targeting OAs.

Table 4 presents the results of the methodological quality assessment using the AMSTAR tool. Seven of the reviews on

---

**Table 1**—References and reasons for article exclusion.

<table>
<thead>
<tr>
<th>Reasons for Exclusion</th>
<th>References of Excluded Reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical guidelines, task force reports, practice parameters</td>
<td>Ngiam et al., 2013; Epstein et al., 2009; Kushida et al., 2006; Thorpy et al., 1995; Verse et al., 2009; Randerath et al., 2011</td>
</tr>
<tr>
<td>Not available</td>
<td>ICRES, 2007</td>
</tr>
<tr>
<td>Included studies with improved AHI only</td>
<td>Bartolucci et al., 2016</td>
</tr>
<tr>
<td>Evaluated predictors of treatment</td>
<td>Alessandri-Bonetti et al., 2015; Okuno et al., 2016; Guarda-Nardini et al., 2015</td>
</tr>
<tr>
<td>Did not assess oral appliances</td>
<td>Manickam et al., 2016; Alsufyani et al., 2013; Veasey et al., 2006; Meoli et al., 2003; Main et al., 2009</td>
</tr>
<tr>
<td>Clinical trial</td>
<td>Furuhashi et al., 2013</td>
</tr>
<tr>
<td>Narrative reviews</td>
<td>Martinez-Gonzalez et al., 2010; Ferguson et al., 2006; Marklund et al., 2012; Schmidt-Nowara et al., 1995; Chan et al., 2007; Sutherland et al., 2011; Hoffstein, 2007; Ng et al., 2005; Fleury et al., 2010; Nimigean et al., 2009</td>
</tr>
<tr>
<td>Short communication</td>
<td>Hoekema et al., 2006</td>
</tr>
<tr>
<td>Commentary</td>
<td>Fox, 2007</td>
</tr>
</tbody>
</table>

---

**Figure 1**—PRISMA flow diagram.
Table 2—Characteristics of included SRs with and without meta-analyses on adult SDB (n = 13).

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Age limits</th>
<th>Numbers of Included Studies</th>
<th>Sample Size</th>
<th>Search end Date</th>
<th>Interventions</th>
<th>Outcomes Measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Ahrens et al., 2011 | Not specified | 14 RCTs (9 Cross-over; 5 Parallel) | 574 | Not reported | • OA vs. control OA  
  • 1-piece vs. another 1-piece OA  
  • 2-piece vs. another 2-piece OA  
  • 1-piece vs. 2-piece OA | AHI or RDI |
| Ahrens et al., 2010 | Not specified | 14 RCTs | 574 | Not reported | • OA vs. control OA  
  • 1-piece vs. another 1-piece OA  
  • 2-piece vs. another 2-piece OA  
  • 1-piece vs. 2-piece OA | Subjective daytime sleepiness: ESS, Fatigue severity scale, sleep disorder questionnaire, sleep symptom questionnaire, HRQoL, Subjective treatment efficacy: non standardized questionnaires or visual analogue scale |

| SRs with Meta-Analyses |            |                             |            |                |               |                  |
| Sharples et al., 2016 (update of McDaid et al., 2009 and Lim et al., 2009) | Adults ≥ 16 years old | 71 RCTs | 6,775 | August 2013 | • OA vs. conservative management  
  • OA vs. CPAP  
  • CPAP vs. conservative management | AHI and ESS |
| Bratton et al., 2015 | Adults ≥ 18 years old | 51 RCTs | 4,888 | August 2015 | • CPAP vs. controls  
  • OAs vs. CPAP  
  • OAs vs. CPAP vs. controls | Change in SBP and DBP |
| Okuno et al., 2014 | Adults ≥ 18 years old | 5 Parallel RCTs | 395 | April 2012 | • OA vs. control OA | AHI, minimum SpO2, arousal index, ESS, Sleep related QoL (using SF-36 Health Survey) |
| Li et al., 2013 | Not specified | 14 RCTs (8 Cross-over; 6 Parallel) | 638 | May 2012 | • OA vs. CPAP | ESS, Health-related QoL, cognitive performance, BP, AHI, arousal index, minimum SpO2, rapid eye movement sleep, treatment usage, treatment preference, side effects, subject preference, and withdrawals |
| Wang et al., 2013 | Not specified | 7 RCTs (5 Crossover; 2 Parallel) | 316 | November 2012 | • OA vs. CPAP | AHI, ESS |
| Iftikhar et al., 2013 | Adults ≥ 18 years old | 7 (2 RCTs; 1 Randomized non-controlled; 1 Non-RCT; 3 Observational) | 399 | December 2011 | • Pre-and post-treatment with OA  
  • OA vs. control OA  
  • OA vs. CPAP vs. Placebo | SBP, DBP, mean arterial BP (MAP), Ambulatory BP monitoring, AHI |
| Health Quality Ontario, 2009 | Adults > 16 years old | 16 RCTs (11 Cross over; 5 Parallel) 5 Systematic reviews | 866 in RCTs | February 2009 | • OA vs. control OA  
  • OA vs. CPAP  
  • OA vs. Surgery  
  • Tongue repositioning devices | AHI, daytime sleepiness, patient preference, compliance, adverse effects |
| McDaid et al., 2009 | Not specified | 48 (6 involved OA) 415 on OA | 1,989 | November 2006 | • CPAP vs. Sham CPAP  
  • CPAP vs. oral placebo  
  • CPAP vs. usual care  
  • OA vs. CPAP  
  • OA vs. CPAP vs. oral placebo  
  • OA vs. CPAP vs. usual care | ESS, MWT, MSLT, BP, cardiovascular events, HRQoL, cognitive function, adverse events, cost-effectiveness (using incremental cost per quality-adjusted life-year). |
| Franklin et al., 2007 | Adults ≥ 19 years old | 30 RCTs (10 involved OA) | 1,850 (587 on OA) | March 2006 | • OA vs. placebo  
  • OA vs. CPAP  
  • OA vs. UPPP  
  • CPAP vs. placebo  
  • Surgery vs. sham or conservative treatment | AHI, ESS, MSLT, BP, MWT, FOSQ, QoL (using SF-36), compliance, side effects |
| Lim et al., 2008 | Adults > 16 years old | 17 RCTs (11 Cross over; 6 Parallel) | 831 | June 2008 | • OA vs. CPAP  
  • OA vs. control OA  
  • OA vs. Surgery | Daytime sleepiness measured by a valid sleep apnea symptom score, QoL, cognitive function, side effects, oxygen desaturation index, one year mortality, patient preference, and BP. |
| Hoekema et al., 2004 | Adults ≥ 21 years old | Studies of efficacy: 16 RCTs (14 Crossover; 2 Parallel; 13 included in meta-analyses) 13 (12 Case series; 1 Case-control) | 499 | 2002 | • OA vs. control OA  
  • Anterior or vertical displacement in OA  
  • Compared different designs of OA  
  • OA vs. UPPP  
  • OA vs. CPAP | Efficacy: AHI [RDI during a full night PSG if AHI not present]  
  • Co-morbidity: Objectively identified side effects of craniofacial/cranio-mandibular complex. |

SR, systematic review; SDB, sleep-disordered breathing; OA, oral appliances; AHI, apnea-hypopnea index; RDI, respiratory disturbance index; CPAP, continuous positive airway pressure; BiPAP, bivelvel positive airway pressure; SpO2, peripheral capillary oxygen saturation; ESS, Epworth Sleepiness Scale; MSLT, Multiple Sleep Latency Test; MWT, maintenance of wakefulness test; FOSQ, functional outcomes of sleep questionnaire; BP, blood pressure; SBP, systolic blood pressure; DBP, diastolic blood pressure; QoL, quality of life; OJ, overjet; OB, overbite; PSG, polysomnography.
Table 3—Characteristics of included SRs with and without meta-analyses on pediatric SDB (n = 4).

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Sample</th>
<th>Numbers of Included Studies</th>
<th>Sample Size</th>
<th>Search End Date</th>
<th>Interventions</th>
<th>Outcomes Measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nazarali et al., 2015</td>
<td>Children &amp; adolescents ≤ 16 years old</td>
<td>4</td>
<td>134</td>
<td>August 2014</td>
<td>• Twin block</td>
<td>• Primary outcome: AHI, RDI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3 Prospective; 1 RCT)</td>
<td></td>
<td></td>
<td>• Herbst appliance with maxillary expander</td>
<td>• Secondary outcomes: oxygen desaturation, daytime and nocturnal symptoms, and dental and skeletal changes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Monobloc with ME screw and tongue retainer vs. no treatment</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Acrylic bite plate vs. no treatment</td>
<td></td>
</tr>
<tr>
<td>Kuhle et al., 2009</td>
<td>Children 1 to 16 years old &amp; studies with up to 20% of children between 17-18 years old</td>
<td>5 Parallel RCTs (1 involved OA)</td>
<td>138</td>
<td>2008</td>
<td>• OA vs. no treatment</td>
<td>• Primary outcomes: AHI or RDI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• ATE vs. Maxillary distraction</td>
<td>• Secondary outcomes: oxygen desaturation index, arousal index, nadir of arterial oxygen saturation, mean SpO2, percentage of children in whom surgical treatment can be avoided, or who have a decrease in tonsillar size, and clinical symptom score (based on parental report).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• CPAP vs. BiPAP</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• TCRF vs. ATE</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Intraoral fluticasone vs. placebo</td>
<td></td>
</tr>
<tr>
<td>Carvalho et al., 2007</td>
<td>Children ≤ 15 years old</td>
<td>1 Quasi-RCT</td>
<td>32</td>
<td>Sept 2005</td>
<td>• OA or functional orthopedic appliances vs. no treatment</td>
<td>• AHI, Reduction of apnea episodes, reduction of upper airway resistance syndrome, reduction of snoring, signs and symptoms of atypical swallowing, speech production disturbance, daytime and nocturnal symptoms, change of mandibular length, improvement of sagittal jaw relation, changes in inter-molar/canine widths, changes in arch perimeter, improvement in OJ and OB, altered growth patters, drop outs and withdrawals, QoL, side effects, economic evaluation, educational outcomes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRs with Meta-Analyses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huynh et al., 2016</td>
<td>Children &amp; adolescents ≤ 18 years old</td>
<td>8 (3 RCTs; 4 Non-RCT; 1 Prospective)</td>
<td>238</td>
<td>April 2014</td>
<td>• Personalized acrylic resin oral plate</td>
<td>• Primary outcomes: AHI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 included in meta-analysis</td>
<td></td>
<td></td>
<td>• Modified monobloc</td>
<td>• Secondary outcomes: oxygen saturation level (%), arousal index, increase of the upper airway volume or structure, sleep quality (%), drop outs and withdrawals.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Endo-oral rapid ME</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Fixed appliances with ME</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• ME, Bi-ME, fixed or removable appliances</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• 2 bands type rapid ME</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Acrylic splint Herbst appliance</td>
<td></td>
</tr>
</tbody>
</table>

SR, systematic review; SDB, sleep-disordered breathing; OA, oral appliances; ATE, adenotonsillectomy; TCRF, temperature controlled radiofrequency ablation; AHI, apnea-hypopnea index; RDI, respiratory disturbance index; CPAP, continuous positive airway pressure; BiPAP, bilevel positive airway pressure; ME, maxillary expansion; SpO2, peripheral capillary oxygen saturation; ESS, Epworth Sleepiness Scale; MSLT, Multiple Sleep Latency Test; MWT, maintenance of wakefulness test; FOSQ, functional outcomes of sleep questionnaire; QoL, quality of life; OJ, overjet; OB, overbite; PSG, polysomnography.

adult OSA were of medium quality\(^5\) and six were of high quality.\(^5\,15\,56,58\) The total mean score for the adult OSA studies was 7.85 ± 1.72 and the scores ranged from 5 to 11. Lie et al.\(^5\) had the highest quality among all the reviews (11/11) and Iftikhar et al.\(^5\) review had the lowest (5/11). Of the reviews on pediatric OSA, three were of high quality\(^5\,15\,61\) and one was medium quality\(^5\) and the total mean score for the studies was 9 ± 1.41 and the range was 7 to 10. Kuhle et al.\(^15\) and Carvalho et al.\(^61\) both had the highest score, whereas Huynh et al.\(^12\) had the lowest score.

The numbers and percentage scores for each response option on the 11 AMSTAR tool domains are presented in Table 5. All reviews had focused questions and a priori designs. The assessment of publication status by searching gray literature sources was conducted in 84.6% of adult reviews and all pediatric reviews. More than 23% of adult reviews failed to include a list or reference the excluded studies, or provide reasons for exclusion. Approximately 85% and 75% of adult and pediatric reviews, respectively, used a valid tool or criteria to assess the quality of included studies. The tools used included one or combination of the following: the American Academy of Sleep Medicine Classification of Evidence, the GRADE approach, the “Risk of bias” tools described in Cochrane Handbook of systematic review of interventions, the Jadad scale, Criteria from Centre for Reviews and Dissemination Report No. 4, and the Schulz et al. tool for allocation concealment. None of the pediatric reviews documented conflicts of interest.

**DISCUSSION**

The aim of this SR was to assess the quality of SRs/MAs about the efficacy of OA in the treatment of adult and pediatric SDB. Twelve SRs with MAs and five without MAs were included. Those five SRs without MAs explained that different definitions...
of OSA treatment success and failure among included studies, varied methodologies, and insufficient evidence were the main reasons for not combining the results quantitatively.

### Adult SDB

All reviews were of medium or high quality according to the AMSTAR criteria. *A priori* design and providing characteristics of included studies are the only two criteria fulfilled by all reviews. The search strategies were well showcased in most reviews. Two reviews,53,54 however, did not report the search dates and four others55,58–60 did not search the gray literature. There was significant variability in the instruments used for quality assessment among the reviews. McDaid et al.51 assessed bias and allocation concealment, but the results were not presented for each study. Hoekema et al.1 examined the efficacy and comorbidity of OAs in the treatment of OSA. For the comorbidity outcomes, the quality of the included studies was assessed based on the general impressions of the investigators and not on valid criteria. Although all the authors of all the reviews documented conflicts of interest, in only one review56 did the authors declare conflicts of interest for the individual studies included within the review.

Meta-analyses were conducted in 11 reviews. Statistical methods were appropriately used to combine the individual studies in most of the reviews (84.6%) except for one.57 In that review,57 randomized trials with both parallel and crossover designs were combined in the MA without justification. Only five reviews (38.5%) either assessed publication bias using funnel plots55,58–60 or described reasons for not investigating it.56 Five of the 11 MAs51,56,58–60 conducted sensitivity analyses to assess the robustness of the results.

The effects of OAs on adult SDB were found to be comparable when high- and moderate-quality adult SRs/MAs were independently evaluated. The six reviews59,52,55,56,59,60 that were identified as having high quality (low risk of bias) found superior efficacy of OAs over controls/placebo/no treatment. These reviews supported a reduction in apnea-hypopnea index (AHI), minimum oxygen saturation level, blood pressure, and arousal index with OAs, in addition to improvement in health-related quality of life. Yet, when compared with CPAP therapy, OAs were of similar or lower efficacy. The other seven moderate quality reviews49,50,52,55,57,58 (moderate risk of bias) reported similar superior efficacy of OAs compared to control/placebo/no treatment. In these reviews, OAs reduced AHI/respiratory disturbance index, improved blood pressure, and were associated with greater patient compliance and satisfaction. Comparisons with CPAP therapy demonstrated similar or superior efficacy of CPAP. Patients however, preferred OAs to CPAP as reported in these SRs.

Most of the primary studies included within the selected SRs were RCTs; however, most of these SRs reported that the quality of the included studies was moderate or low. Blinding and allocation concealment were impossible in many individual studies due to the differences between the interventions (i.e., OA vs. CPAP). Also, intent-to-treat analyses were not conducted in many studies, which may have introduced bias. There seems to be a dire need for better RCTs about the use of OAs in patients with SDB. There is also a need for a standard definition of SDB/OSA treatment success. The definitions of treatment success used in most trials were varied including 50% reduction in AHI, final AHI of less than 5, 10, or 15, and improvement in subjective symptoms. Treatment success

---

### Table 4—Quality assessment of included reviews using the AMSTAR criteria.

<table>
<thead>
<tr>
<th>Author(s), year of publication</th>
<th>1 A priori design</th>
<th>2 Duplicate study selection and data extraction</th>
<th>3 Comprehensive literature search</th>
<th>4 Publication status as an inclusion criteria</th>
<th>5 Provided list of included and excluded studies</th>
<th>6 Provided characteristics of included studies</th>
<th>7 Assessed quality</th>
<th>8 Appropriate use of quality to form conclusions</th>
<th>9 Appropriate use of methods to combine findings</th>
<th>10 Assessed publication bias</th>
<th>11 Documented conflict of interest</th>
<th>Total score (/11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult SDB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharples et al., 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>9</td>
</tr>
<tr>
<td>Bratton et al., 2015</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>9</td>
</tr>
<tr>
<td>Okuno et al., 2014</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>9</td>
</tr>
<tr>
<td>Wang et al., 2013</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>6</td>
</tr>
<tr>
<td>Iftikhar et al., 2013</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Li et al., 2013</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Ahrens et al., 2011</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>N</td>
<td>8</td>
</tr>
<tr>
<td>Ahrens et al., 2010</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>N</td>
<td>8</td>
</tr>
<tr>
<td>McDaid et al., 2009</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>8</td>
</tr>
<tr>
<td>Health Quality Ontario, 2009</td>
<td>Y</td>
<td>CA</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>6</td>
</tr>
<tr>
<td>Franklin et al., 2007</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>9</td>
</tr>
<tr>
<td>Lim et al., 2006</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>9</td>
</tr>
<tr>
<td>Hoekema et al., 2004</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>7</td>
<td>N</td>
<td>7</td>
</tr>
<tr>
<td>Pediatric SDB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hugm et al., 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>7</td>
</tr>
<tr>
<td>Nacacoli et al., 2015</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>9</td>
</tr>
<tr>
<td>Kahle et al., 2009</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Canavoli, et al., 2007</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Scoring choices are “Y”: Yes, “N”: No; “CA”: Can’t answer, and “NA”: Not applicable. “Y” items received one point and the sum was calculated. Reviews were considered high quality if scored (9–11), medium or low if scored (5–8) and (< 5) respectively.
depends on the severity of the disease and the design and degree of mandibular forward repositioning of OAs among other factors. Therefore, the development of a standard objective definition of treatment success and a further understanding of the predictors of treatment success are warranted in future studies.

The duration of follow-up also varied greatly between the trials. Trials of shorter duration might have exaggerated the treatment efficacy as compared to trials of longer durations. Long-term stability of the changes was not assessed.

**Pediatric SDB**

Very few reviews were identified and included in this study. In one, Carvalho et al.\(^6\) reported findings from only one quasi-randomized trial with high risk of bias. Only two of the four reviews (50%) assessed publication bias.\(^5,6\) Similar to studies on adults, none of the reviews reported conflicts of interest among the primary studies. The small number of reviews on pediatric OSA reflects lack of trials on the controversial effects of the different modalities to treat OSA or persistent OSA after adenoidectomy and/or tonsillectomy in children and adolescents. Due to the great heterogeneity of the studies, only one SR\(^12\) conducted a MA. Even there, the applicability of a MA is questionable due to the methodological heterogeneity between the included studies. Future studies are warranted about this important subgroup that seems to behave differently than the adult subgroup. In addition, normal or altered craniofacial growth changes come into play in this subsample. The effect of treatment on growth and the subsequent prognosis/risk for OSA has to be considered. For example, orthodontic treatment modalities that involve restriction of maxillary growth may worsen disease prognosis.

When the outcomes of OA therapy were compared between high- (low risk of bias) and moderate-quality (moderate risk of bias) SRs, only one high-quality SR\(^5\) reported significant improvement in AHI after OAs, but with no cure or normalization of AHI. However, the other two high-quality SRs\(^6,16\) and the one moderate-quality SR\(^12\) all reported insufficient evidence of OAs efficacy due to limited number of studies. Therefore, evidence on the efficacy of OA therapy for pediatric OSA yielded similar conclusions irrespective of the degree of risk of bias (significant consistency among studies).

**Limitations**

Similar to PRISMA, which only assesses the quality of reporting in the SRs, and not the risk of bias among the included studies inside those SRs, high AMSTAR scores should not be

<table>
<thead>
<tr>
<th>AMSTAR criteria</th>
<th>Sample</th>
<th>“Yes” n (%)</th>
<th>“No” n (%)</th>
<th>“Can’t answer” n (%)</th>
<th>“Not applicable” n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A priori design</td>
<td>Adult</td>
<td>13 (100)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>Pediatric</td>
<td>4 (100)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>2. Duplicate study selection and data extraction</td>
<td>Adult</td>
<td>12 (92.3)</td>
<td>0 (0.0)</td>
<td>1 (7.7)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>Pediatric</td>
<td>4 (100)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>3. Comprehensive literature search</td>
<td>Adult</td>
<td>12 (92.3)</td>
<td>1 (7.7)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>Pediatric</td>
<td>4 (100)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>4. Publication status as an inclusion criteria</td>
<td>Adult</td>
<td>11 (84.6)</td>
<td>2 (15.3)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>Pediatric</td>
<td>4 (100)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>5. Provided list of included and excluded studies</td>
<td>Adult</td>
<td>3 (23.1)</td>
<td>10 (76.9)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>Pediatric</td>
<td>4 (100)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>6. Provided characteristics of included studies</td>
<td>Adult</td>
<td>13 (100)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>Pediatric</td>
<td>4 (100)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>7. Assessed quality</td>
<td>Adult</td>
<td>11 (84.6)</td>
<td>2 (15.4)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>Pediatric</td>
<td>3 (75.0)</td>
<td>1 (25.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>8. Appropriate use of quality to form conclusions</td>
<td>Adult</td>
<td>9 (69.2)</td>
<td>4 (30.8)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>Pediatric</td>
<td>3 (75.0)</td>
<td>1 (25.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>9. Appropriate use of methods to combine findings</td>
<td>Adult</td>
<td>11 (84.6)</td>
<td>2 (15.4)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>Pediatric</td>
<td>4 (100)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>10. Assessed publication bias</td>
<td>Adult</td>
<td>5 (38.5)</td>
<td>6 (46.2)</td>
<td>0 (0.0)</td>
<td>2 (15.4)</td>
</tr>
<tr>
<td></td>
<td>Pediatric</td>
<td>2 (50.0)</td>
<td>1 (25.0)</td>
<td>0 (0.0)</td>
<td>1 (25.0)</td>
</tr>
<tr>
<td>11. Documented conflict of interest</td>
<td>Adult</td>
<td>1 (7.6)</td>
<td>12 (92.3)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>Pediatric</td>
<td>0 (0.0)</td>
<td>4 (100)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
</tbody>
</table>

**Table 5—AMSTAR response rates by criterion.**
CONCLUSIONS

SRs/MAs on the efficacy of OAs in the management of SDB/OSA are in general terms of acceptable quality, but there is still room for improvement. SRs/MAs on both adult and pediatric SDB/OSA can be strengthened by assessing publication bias in the meta-analyses and by documenting conflict of interest of the primary studies included within the SRs/MAs.

Overall, OAs showed superior efficacy to no treatment/controls in the management of adult SDB/OSA, but evidence is insufficient to support the efficacy of those appliances in pediatric SDB/OSA.

ABBREVIATIONS

AMSTAR, Assessment of Multiple Systematic Reviews
MA, meta-analyses
OA, oral appliances
OSA, obstructive sleep apnea
RCTs, randomized controlled trials
SDB, sleep-disordered breathing
SR, systematic review

REFERENCES


SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication February, 2016
Submitted in final revised form May, 2016
Accepted for publication June, 2016

Address correspondence to: Thikriat Al-Jewair, BDS, MBA, MSc, FRCD(C), Department of Orthodontics, State University of New York at Buffalo, NY; Tel: (716) 713-4547; Email: thikriat@buffalo.edu

DISCLOSURE STATEMENT

This was not an industry supported study. Carlos Flores-Mir is a speaker for Opal Orthodontics regarding Class II malocclusion management in pediatric populations. The other authors have indicated no financial conflicts of interest.