SYSTEMATIC REVIEW AND META-ANALYSIS O

Systematic Review and Meta-Analysis of Passive Lower Lingual Arch for Resolving Mandibular Incisor Crowding and Effects on Arch Dimension

Chia-Yu Chen, DDS[1] • Kuei-Ling Christine Hsu, DDS, MS[2] • Abdullah A. Marghalani, BDS, MSD, DrPH[3] • Vineet Dhar, BDS, MDS, PhD[4] • James A. Coll, DMD, MS[5]

Abstract: ***Purpose:*** *This systematic review and meta-analysis assessed whether passive lower lingual arch (LLA) resolves mandibular incisor crowding and affects mandibular arch dimension.* ***Methods:*** *We searched PubMed, Web of Science, and Cochrane Database of Systemic Reviews for both randomized controlled trials and nonrandomized studies from 1940 to March 2018. Inclusion criteria were healthy children in mixed dentition with mandibular incisor crowding treated with LLA. Our primary outcome was the amount of mandibular incisor crowding resolved after LLA, and secondary outcomes were effects of LLA mandibular arch dimension changes versus untreated controls (UTCs).* ***Results:*** *From 559 screened articles, seven qualified for systematic review and meta-analyses. The average resolution of mandibular incisor crowding after LLA was 5.10 mm (P=.001) with the evidence assessed as very low quality. Arch perimeter and arch length changes were not significantly different between LLA and UTCs (P=0.20 and P=0.87, respectively). There were significant small increases of 0.79 mm in intercanine width (P<.001) and 0.69 mm in intermolar width (P=.003) with a low and a very low quality of evidence, respectively.* ***Conclusions:*** *Lower lingual arch was effective in resolving mandibular incisor crowding without any significant arch perimeter or arch length changes of greater than one mm.* *(Pediatr Dent 2019;41(1):9-19.E1-E3) Received September 5, 2018 | Last Revision December 7, 2018 | Accepted December 15, 2018*

KEYWORDS: TOOTH CROWDING, META-ANALYSIS, INTERCEPTIVE ORTHODONTICS

In the mixed dentition, mandibular incisor crowding is a common occurrence that may be resolved using the leeway space.[1,2] According to Nance,[3] the total mandibular leeway space is 3.4 mm. However, this value can vary between subjects and studies.[4] The size difference is between the mesial-distal width of the primary mandibular second molar and the second premolar and is often referred to as 'E' space.[5] Preservation of the available leeway space may provide the space needed to resolve the mandibular incisor crowding.

A passive lower lingual arch (**LLA**) has been widely used in interceptive orthodontics to preserve the leeway space and maintain the arch length.[6] In the mixed dentition, the LLA can maintain the 'E' space after premature loss of the primary second molars but also preserve the leeway space in crowded cases to avoid premolar extractions and resolve mandibular incisor crowding.[7-13] The effect of LLA on arch dimensions has also been reported. Increases in intercanine and intermolar width after LLA placement were consistently reported in different observational cohort studies, ranging from 0.72 mm to 2.41 mm.[7,8,13] Despite the effectiveness of the LLA resolving mandibular incisor crowding, Brennan et al.[13] and De Baets et al.[8] reported that there was a significant decrease in arch perimeter, measured from the combined distance from the mesial contact point of mandibular central incisors to the mesial contact points of the first permanent molars. The authors attributed this decrease in arch perimeter to lingual tipping of the incisors.

The management of leeway space is more critical in the mandible than in the maxilla, since there is limited capability for expansion and labial movement of incisors can be unstable.[16] Numerous case reports used an LLA as a treatment strategy to deal with a child's malocclusion.[17-19] However, there is only one published systematic review reporting on the effect of LLA on arch dimensions in the developing dentition.[20]

Further investigation is needed to determine the amount of mandibular incisor crowding the LLA can resolve and its effect on three arch dimensions: (1) arch perimeter; (2) arch length; and (3) arch width. The purposes of this systematic review and meta-analysis were to evaluate, in healthy children in the mixed dentition having mandibular incisor crowding: (1) how many millimeters of mandibular incisor crowding can be resolved by a lower lingual arch; and (2) the effect of using an LLA compared to not using one on arch dimensions.

Methods

Search strategy and information sources. To identify all the studies regarding the effectiveness of LLA in resolving mandibular incisor crowding, PubMed, Web of Science, Cochrane Database of Systemic Reviews, and the grey literature through OpenGrey were searched for articles published between January 1940 and March 2018. The concept of leeway space was first established in the 1940s by Nance,[3] and the utilization of LLA for space management was later discovered. Therefore, the year 1940 was set as the starting date for our search. The following key terms were used in the literature search, as Medical Subject Headings terms or free text words, joined by "or" and "and": "lingual", "nance", "arch", "arches", and "orthodontics" or "orthodont*" (see **Electronic Appendix: Section 1** for search strategy). No restriction on language was placed; however, all eligible studies were in English. Relevant journals

[1]Dr. Chen is a clinical instructor, Drs. [4]Dhar and [5]Coll are clinical professors, and [2]Dr. Hsu is a clinical assistant professor, all in the Division of Pediatric Dentistry, School of Dentistry, University of Maryland, Baltimore, Md., USA. [3]Dr. Marghalani is a faculty member, Department of Preventive Dentistry, Umm Al-Qura University Faculty of Dentistry, Makkah, Saudi Arabia.
Correspond with Dr. Coll at jimcolldmd@gmail.com

Supplemental material available in the online version.

and references listed in one systematic review were also hand-searched. The search was performed by one author and a librarian at the Health Sciences and Human Services Library, Baltimore, Md., USA.

After removing the repeated studies, title and abstract and full-text review was performed by two reviewers to identify studies for inclusion or exclusion in this systematic review. If the abstract of a study did not provide enough information, a full text was accessed for further review. If a study was excluded after full-text review, the reasons for exclusions were described. If the two screening reviewers could not agree on selected studies, a third person would be contacted to resolve any disagreement; however, there were no unresolved issues requiring a third party.

Eligibility criteria. The population, intervention, comparison, outcomes and study design (**PICO-S**) method was used to establish the inclusion and exclusion criteria. The population was defined as healthy children in mixed dentition with mandibular incisor crowding. The intervention was defined as placement of LLA on permanent mandibular first molars (**L6s**). The comparison was to no intervention in the mandible. The outcomes measures are described later, but an eligible study must have reported satisfactory statistical data for inclusion. Since there were only two randomized controlled trials (**RCTs**) on LLAs after 1980,[9,14] we included nonrandomized studies (**NRSs**), such as observational analytic studies (including prospective/ retrospective cohort studies, case control studies, and cross-sectional studies), and nonrandomized experimental clinical trials, in addition to RCTs. Systematic reviews (**SRs**) were also included in our search for studies regarding LLA and mandibular dimensions.

The following exclusion criteria were used to reduce the impact of confounders or unpredictable contributing factors. Studies in which the participants had reported skeletal discrepancy or deficiency were excluded. Also, studies were excluded in which the treatment involved more than passive LLA use, such as mandibular arch expansion appliances, mandibular molar distalization appliances, and premolar extractions. Furthermore, studies that involved maxillary appliances, such as headgear or two-by-four appliances, were not excluded because prior research had indicated that there was no significant effect of a maxillary appliance on the position of the mandibular molars.[10] To perform a comprehensive review, we didn't exclude studies in which different gauges of arch wire, omega loops, or removable-design of LLA was used if such appliances were passively used for space maintenance. These factors were recorded for possible sensitivity analysis.

Outcomes. The primary outcome was the changes in mandibular incisor crowding in millimeters (mm) after the placement of LLA. Crowding was described as tooth size/ arch size discrepancy (**TSASD**)[13] or Little's irregularity index (**LII**).[21] The difference between the combined mesial-distal tooth size of permanent teeth and the available arch space is termed TSASD. When the combined mesial-distal tooth size exceeds the available arch space, crowding occurs and TSASD is identified as a negative value. LII measures the horizontal linear displacement of anatomic contact points of each mandibular incisor from the adjacent anatomic point and sums the five displacements. The positive value indicates the degree of anterior irregularity representing mandibular incisor crowding. Both methods of crowding were recorded in this systematic review.

The secondary outcomes were changes in arch perimeter, arch length, intercanine width, and intermolar width. Arch perimeter was defined as the combined distance between the mesial anatomic contact points of bilateral L6s to the contact point between the mandibular central incisors (**L1s**), which other authors define as arch length[7-9,13-15] or incisor-to-molar distance.[10] Arch length was defined as the distance from a point bisecting the mesial anatomic contact points of L6s to the contact point of the L1s, which other authors defined as arch depth.[9,14] Intercanine width was defined as the distance between the cusp tips or estimated cusp tips if wear facets showed on bilateral mandibular canines. Intermolar width was defined as the distance between a fixed reference point on the L6s. Different reference points had been reported, such as the central fossa and the mesial-buccal cusp tips. In our meta-analysis, we focused on the changes of intermolar width over the treatment/observation period, regardless of the choice of reference points.

Data extraction. Three authors performed data extraction independently. A standardized form was generated by two authors, in which 36 fields were included for data extraction (e.g., author name, publication year, PICO-S components). Not all the studies reported every assessed outcome, so data in these studies was extracted in as many of the 36 fields as possible. Disagreements were resolved through discussion by the three authors. For studies that had only examined/reported one group of participants treated with LLA, the data was extracted to compare the effect from before LLA and after LLA was in place for a specified period of time. For the studies with more than two intervention groups, only the intervention group comparing an LLA group to an untreated control (**UTC**) were considered.

Data synthesis. Meta-analysis was performed using a random effect model to estimate the effect size, mean difference (**MD**), and 95 percent confidence interval (**95% CI**) for each outcome measure. For paired data, which were the prepost studies, the correlation was calculated, if not reported, and the weighted correlation mean was assumed for studies that did not report enough information to estimate the correlation (standard deviations of the baseline, postintervention, and difference). We attempted to estimate the standard deviation when it was missing.[22] Heterogeneity was evaluated using the I^2 statistic and the Cochrane test for heterogeneity, with a P-value less than 0.1 considered to be statistically significant. The analyses were performed using Comprehensive Meta-Analysis (CMA) 3.0 software (Biostat, Englewook, N.J., USA).

Risk-of-bias (ROB) assessment. The quality of the included studies was evaluated independently using the revised Cochrane Collaboration's risk of bias assessment tool for RCTs (**RoB 2.0**)[23] and Risk Of Bias In Non-Randomized Studies of Interventions (**ROBINS-I**)[24] for NRSs by three of the authors.

The overall assessment of ROB for RCTs was based on five key domains using RoB 2.0, including ROB arising from the randomization process, bias due to deviations from intended interventions, bias in measurement of the outcome, bias due to missing outcome data, and bias in selection of the reported result. After evaluating each key domain, we determined the overall ROB assessment for each study, as follows: "low" ROB if all domains were low ROB; "some concern" if at least one domain was judged to have some concern; and "high" ROB if at least one domain had high ROB or multiple domains were judged to have some concerns in a way that substantially lowered confidence in the result.

For NRSs, the overall assessment of ROB was based on seven key domains using ROBINS-I, including bias due to confounding, bias in selection of participants into the study, bias in classification of interventions, bias due to deviations from intended interventions, bias due to missing data, bias in measurement of outcomes, and bias in selection of the reported results. After evaluating each key domain, we determined the overall ROB for each study as followed: low ROB when all domains were at low ROB; moderate ROB when all domains were at either low or moderate ROB; serious ROB when at least one domain was judged as serious ROB; critical ROB when at least one domain was at critical ROB; and no information when there was a lack of information in one or more key domains.

Grading. The quality of evidence of each outcome in the meta-analysis was evaluated using the Grading of Recommendations Assessment, Development, and Evaluation (**GRADE**) system. The following criteria were included for assessment of the quality of evidence for each outcome across studies: study design; ROB; consistency; precision; publication bias; and other considerations. In evaluating study design, experimental studies provided higher quality evidence when compared to observational studies in general. ROB was determined by using the ROB tools (RoB 2.0 and ROBINS-I) and GRADE guidance tool. Consistency was judged based on the heterogeneity (I^2) of each outcome, and was ranked as: not serious—zero to 30 percent; serious—30 to 75 percent; and very serious—greater than 75 percent. Precision was judged based on the crossing of the CI of the pooled outcome to the no-effect line and the total sample size; it was ranked as "not serious" if total sample size was larger than 40, "serious" if between 20 and 40, and "very serious" if smaller than 20. Publication bias was assessed when outcomes had more than 10 articles included for quantitative analysis. Other considerations, including large magnitude of effect and plausible confounders, were also accessed for observational studies. The GRADE system results in four grades in rating the quality of evidence: (1) high; (2) moderate; (3) low; and (4) very low (see **Electronic Appendix: Section 2** for GRADE assessment).

Results

Description of studies. The process of evaluating articles for inclusion in this study is illustrated in Figure 1, based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart. The search strategy yielded a total of 559 nonduplicate reports from all databases. After title screening, 542 records were excluded with reason. Of these, there were 483 articles excluded because the titles revealed irrelevant study scope. Abstracts from 76 were reviewed, and 59 more were excluded due to wrong study design, leaving 17 articles for full-text review. After full-text review, 10 additional articles were excluded due to the following reasons: four studies[25-28] included wrong interventions involving mandibular arch expansion, premolar extractions or other kinds of space maintainer; one article[29] was excluded due to the wrong population (skeletal class II participants); three articles had the wrong outcome,[30-32] focusing on the side effects following the placement of the LLA; and two articles[33,34] had wrong outcomes studying the effect of LLA on positional changes of permanent teeth. Finally, seven studies met the inclusion criteria and qualified for one or more meta-analyses.

The demographic characteristics of the seven included studies are summarized in Table 1. This systematic review included four retrospective cohort studies,[7,8,13,15] one experimental NRS,[10] and two RCTs[9,14] with a total of 307 participants treated with LLA and 74 participants served as UTCs. All the participants were reported to be healthy without any facial deformity or skeletal discrepancies. One study focused on Caucasian population,[15] while the other studies did not mention ethnicity. The age of the participants at the initial evaluation ranged from seven to 13 years old. The treatment/observation period ranged from 10.5 months to four years. One study did not mention if the operator was a specialist,[10] whereas the remaining trials were conducted either in private orthodontic clinics or orthodontic departments of a university by orthodontists. The reasons for placing LLA were to resolve incisor crowding (three studies[7,8,13]) and to save E space (four studies[9,10,14,15]). When primary molar extraction was involved in the treatment plan, the extraction was performed before the insertion of LLA in one study[13] and after insertion in four studies,[8-10,14] and there was no mention of the timing of extraction in the remaining two studies.[7,15] Five studies used only LLA,[8,9,13-15] whereas the remaining two studies involved maxillary headgear or two-by-four appliances in the treatment plan.[7,10] The LLAs used in the studies were all passively inserted with removable design in two studies[7,10] and with fixed design in five studies.[8,9,13-15] The LLA was constructed of 0.9 mm (0.036 inch) stainless steel wire (**SSW**) in four studies,[9,10,13,15] one of which compared the effects of 0.9 mm SSW to 1.25 mm (0.050 inch) SSW,[9] 0.8 mm (0.032 inch) SSW in one study,[14] and 0.76 mm (0.030 inch) SSW in one study[7]; one study didn't report the gauges of the arch wire.[8]

Quality assessment. Table 2 presents the individual domain ROB based on RoB 2.0 for the two RCTs.[9,14] The randomization sequence generation process in Owais' study[9] was not clearly addressed. Rebellato's study[14] provided no information on how subjects were randomized. Both studies

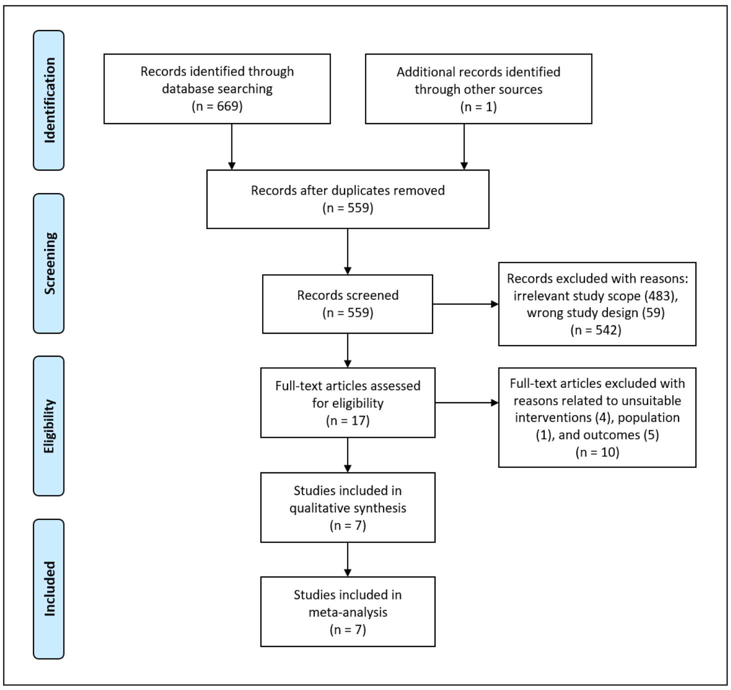

Figure 1. Flowchart of study selection process.

failed to fulfill allocation concealment in the trials; therefore, both were deemed having some concerns in randomization process. In the study by Owais et al., there was no reporting on blinding of the assessors; therefore, some concerns were placed for measurement of this outcome. No clear information regarding missing outcome data was given in both RCTs. The overall ROB assessments for both RCTs was determined as "some concerns."

Table 3 presents the ROB assessment based on ROBINS-I for the five NRSs.[7,8,10,13,15] Regarding confounding, Singer's study was ranked as moderate ROB because there was a baseline age discrepancy between LLA group and UTCs. Regarding the selection of participants into the study, the study by Dugoni et al.[7] was ranked as serious ROB because the selection criteria was related to the post-treatment outcome. In the study by De Baets et al.,[8] the treatment groups were divided based on the patients' final appearance following the placement of LLA; however, it was adjustable during quantitative analysis—thus, it was ranked as moderate ROB.

Table 1. DEMOGRAPHIC CHARACTERISTICS SUMMARY OF THE INCLUDED STUDIES*

Author (year), country	Study design	LLA group	UTC group	Mean age after LLA/study duration
Brennan and Gianelly (2000)[13], USA	Observational, retrospective cohort	N=107 Mean age (years)†: 8.6 (7-11 yrs) All with fixed LLA Wire gauge: 0.9 mm	Not reported	Not reported
Dugoni et al. (1995)[7], USA	Observational, retrospective cohort	N=25 Mean age (years)†: 8.2 (7-11 yrs) All with removable LLA and maxillary 2x4 appliance Wire gauge: 0.76 mm	Not reported	Mean age (years)‡: 13.5
De Baets et al. (1995)[8], Switzerland	Observational, retrospective cohort	N=39 Age range (years)†: 7-13 yrs All with fixed LLA and/or maxillary interceptive orthodontics	Not reported	Study duration (years): 4 yrs
Fichera et al. (2011)[15], Italy	Observational, retrospective cohort	N=42 Mean age (years)†: 9.0 yrs All with fixed LLA Wire gauge: 0.9 mm	N=18 Mean age (years)†: 9.2 yrs	Mean age (years)‡: LLA: 12.0 UTC: 11.8
Owais et al. (2011)[9], Jordan	Experimental, RCT	N=44 Mean age (years)†: 10.6 yrs Wire gauge: 0.9 mm (N=20), 1.25 mm (N=24)	N=23 Mean age (years)†: 10.6 yrs	Mean age (years)‡: LLA: 1.34 UTC: 2.42
Rebellato et al. (1997)[14], Israel	Experimental, RCT	N=14 Mean age (years)†: 11.5 yrs Wire gauge: 0.8 mm	N=16 Mean age (years)†: 11.3 yrs	Mean age (months)‡: LLA: 10.5 UTC: 12.5
Singer et al. (1974)[10], USA	Experimental, NRS	N=36 Mean age (years)†: 11.3 yrs All with removable LLA and/or maxillary headgear Wire gauge: 0.9 mm	N=17 Mean age (years)†: 10.2 yrs No LLA, with or without maxillary headgear	Study duration (years): LLA: 0.9 UTC: 1.8

* RCT=randomized clinical trial; NRS=nonrandomized study; LLA=lower lingual arch group; UTC=untreated control. † Age at initial examination. ‡ Age after treatment.

Table 2. RISK OF BIAS (ROB) ASSESSMENT FOR RANDOMIZED CONTROLLED TRIALS BASED ON COCHRANE ROB 2.0

Author, year	Risk of bias arising from:					Overall ROB assessment
	Randomization process	Deviations from intended interventions	Measurement of the outcome	Missing outcome data	Selection of the reported result	
Owais et al., (2011)[9]	Some concerns	Low risk	Some concerns	Some concerns	Low risk	Some concerns
Rebellato et al., (1997)[14]	Some concerns	Low risk	Low risk	Some concerns	Low risk	Some concerns

Regarding the measurement of outcomes, the studies by Brennan and Gianelly[13] and De Baets et al.[8] were ranked as moderate ROB because no information was given regarding if the examiners were blinded. The study by Fichera et al.[15] was ranked as serious ROB because the outcome assessor and the treatment performer were the same person and no further information was given if the assessor was blinded during measuring outcome. Regarding missing data and selection of the reported results, Singer's study[10] was ranked as moderate ROB, as there were nine missing data on intercanine width measurements without proper explanation. In addition, they didn't report the outcomes on mandibular dimensional changes in UTCs as well as the standard deviation of the LLA group. Thus, the overall ROB was ranked as moderate ROB in three studies[8,10,13] and as serious ROB in two studies,[7,15] based on ROBINS-I. Regardless of the results of ROB assessment, all selected studies were enrolled in one or more meta-analysis with a GRADE evaluation for each outcome (see **Electronic Appendix: Section 2** for GRADE assessment).

Meta-analysis results on resolution of mandibular incisor crowding comparing pre- and postplacement of LLA. Three retrospective cohort studies[7,8,13] evaluated the effectiveness of the LLA in resolving mandibular incisor crowding before and after placement in patients seven to 13 years of age. The study by Brennan and Gianelly[13] used TSASD for measuring mandibular incisor crowding, while the studies by Dugoni et al. and De Baets et al.[7,8] used LII for measurement. The study by Brennan and Gianelly showed a significant resolution of 5.0±2.1 mm of mandibular incisor crowding after the placement of LLA; however, it could not be used in the meta-analysis[35] due to them using a different measurement (TSASD). The meta-analysis synthesized from the two studies,[7,8] using the same measurement tool (LII), showed a statistically significant decrease of mandibular incisor crowding after the placement of LLA, irrespective of the wire gauge and appliance design (N equals 59; MD equals 5.10 mm; 95% CI equals 2.15 to 8.05; P=.001; Figure 2). The quality of evidence was assessed as very low, based on them being observational study designs with no comparator, moderate to serious ROB, and high heterogeneity (I^2 equals 87.24 percent).

Meta-analysis results on mandibular arch dimensional changes comparing LLA group to UTCs: *arch perimeter.* Two RCTs[9,14] and one retrospective cohort study[15] investigated the effect LLA had in maintaining arch perimeter. In the RCT by Owais et al.,[9] there were two experimental groups using two different SSW thicknesses (0.9 mm and 1.25 mm); therefore, the groups were computed as two separate arms in the meta-analysis. The meta-analysis on the changes in arch perimeter showed that there was no significant difference between the LLA group (N equals 100) and UTCs (N equals 80) over the treatment period (MD equals 0.97 mm; 95% CI equals -0.50 to 2.44; P=0.20; Figure 3). The quality of evidence was assessed as very low due to high heterogeneity (I^2 equals 99.42 percent) and few studies.

Table 3. RISK OF BIAS (ROB) ASSESSMENT FOR NONRANDOMIZED STUDIES BASED ON RISK OF BIAS IN NONRANDOMIZED STUDIES OF INTERVENTIONS (ROBINS-I)*

Author, year	Confounding	Selection of participants into the study	Classification of interventions	Deviations from intended interventions	Missing data	Measurement of outcomes	Selection of the reported results	Overall ROB assessment
Brennan and Gianelly, (2000)[13]	Low	Low	Low	Low	Low	Moderate	Low	Moderate
Dugoni et al., (1995)[7]	Low	Serious	Low	Low	Low	Low	Low	Serious
De Baets et al., (1995)[8]	Low	Moderate	Low	Low	Low	Moderate	Low	Moderate
Fichera et al., (2011)[15]	Low	Low	Low	Low	Low	Serious	Low	Serious
Singer et al., (1974)[10]	Moderate	Low	Low	Low	Moderate	Low	Moderate	Moderate

* ROB ranking based on RoB 2.0 for RCTs=low risk, some concerns, and high risk; ranking of ROB based on ROBINS-I=low, moderate, serious, critical, and no information.

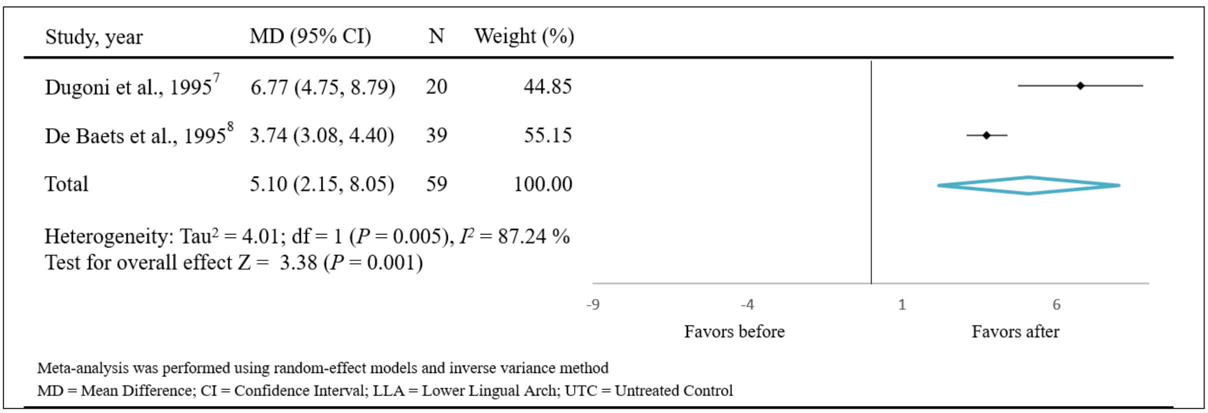

Figure 2. Forest plot for the resolution of mandibular incisor crowding comparing pre- and postplacement of lower lingual arch (before and after LLA).

Arch length. Two RCTs[9,14] evaluated the effect of LLA in maintaining arch length. The meta-analysis on the changes in arch length showed that there was no significant difference between LLA group (N equals 58) and UTCs (N equals 62) over the treatment period (MD equals 0.09 mm; 95% CI equals -0.93 to 1.10; *P*=0.87; Figure 3). The quality of evidence was assessed as low due to high heterogeneity (I^2 equals 82.37 percent).

Intercanine width. One retrospective cohort study[15] and one RCT[9] evaluated the changes of intercanine width between groups. Different gauges of SSW, 0.9 mm and 1.25 mm, were used and evaluated by Owais et al.,[9] and the data from two

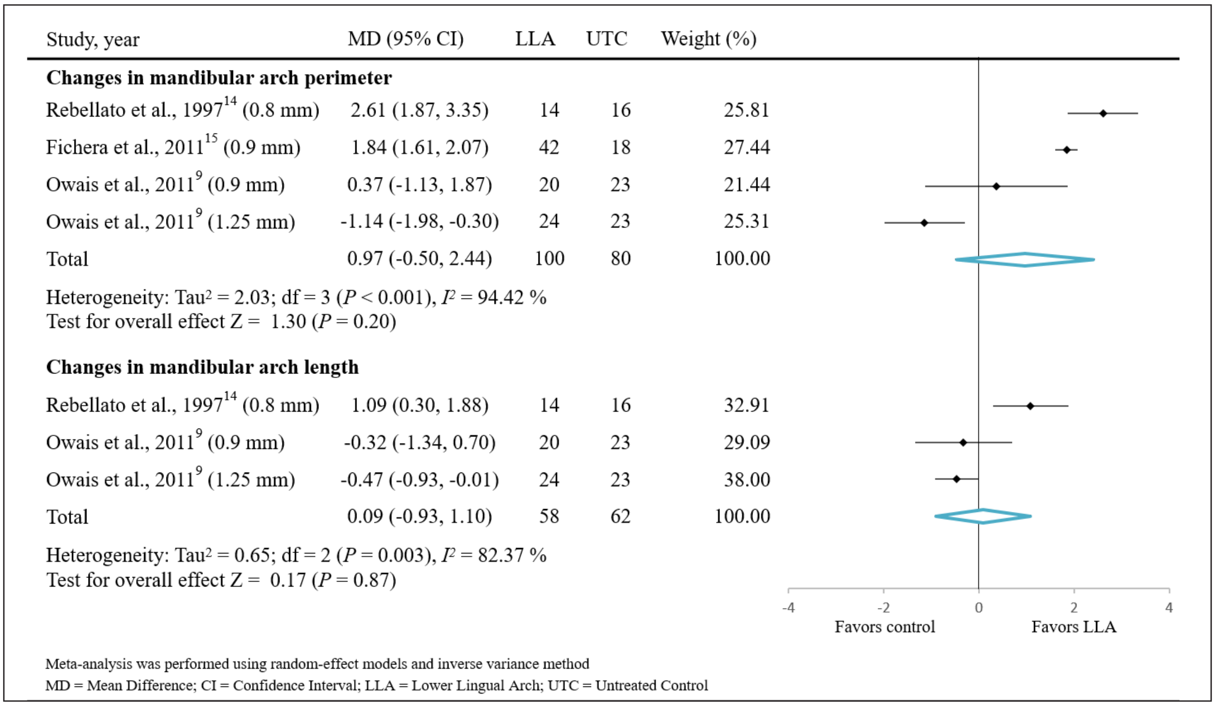

Figure 3. Forest plots for the changes of mandibular arch perimeter and arch length for lower lingual arch group and control group (untreated controls).

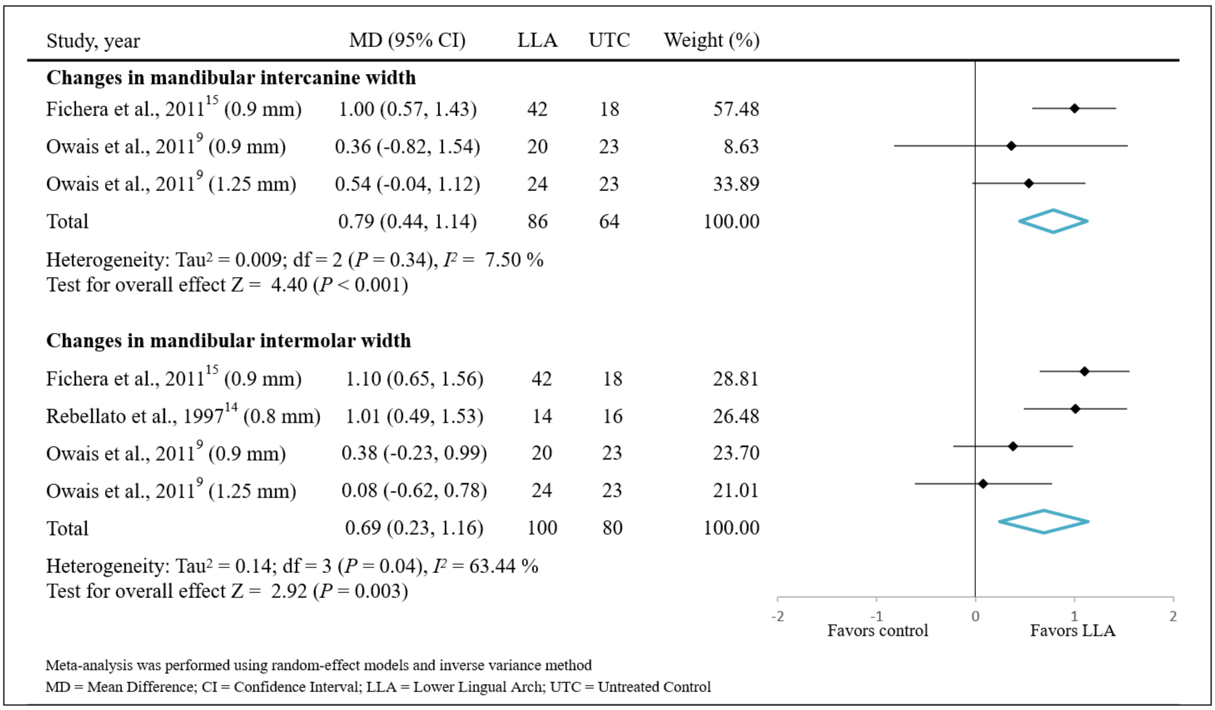

Figure 4. Forest plots for the changes of mandibular arch width for LLA group and control group (untreated controls).

arms were computed into the meta-analysis separately. The meta-analysis showed that there was a significant increase in intercanine arch width in LLA group (N equals 86) compared to that in UTCs (N equals 64) over the treatment period (MD equals 0.79 mm; 95% CI equals 0.44 to 1.14; *P*<.001; Figure 4). The quality of evidence was assessed as low due to diverse study designs and serious ROB.

Intermolar width. One retrospective cohort study[15] and two RCTs[9,14] evaluated the changes of intermolar width between groups. The reference points of L6s for this outcome measure differed, since the central fossa was used in two studies[9,15] and the mesial-buccal cusp tips were used in the other.[14] Different gauges of SSW, 0.9 mm and 1.25 mm, were used and evaluated by Owais et al.,[9] and the data from the two arms were computed into the meta-analysis separately. The meta-analysis showed that there was a significant increase in intermolar arch width in the LLA group (N equals 100) compared to that for UTCs (N equals 80) over the treatment period (MD equals 0.69 mm; 95% CI equals 0.23 to 1.16; *P*=.003; Figure 4). The quality of evidence was assessed as very low due to diverse study designs and high heterogeneity (I^2 equals 63.44 percent).

Sensitivity analysis. In the RCT by Owais et al.,[9] there were two experimental groups using two different thicknesses of SSWs (0.9 mm and 1.25 mm). When performing the meta-analysis, we found that the gauge of the SSW had a different impact on arch dimension changes, especially on arch perimeter. Therefore, the experimental groups in Owais' study[9] were computed separately to assess the effect of different gauges of SSW in the meta-analyses. Intermolar width was measured by Rebellato et al.[14] using the medial buccal cusp tips of L6s. The other authors[9,15] used the L6s central

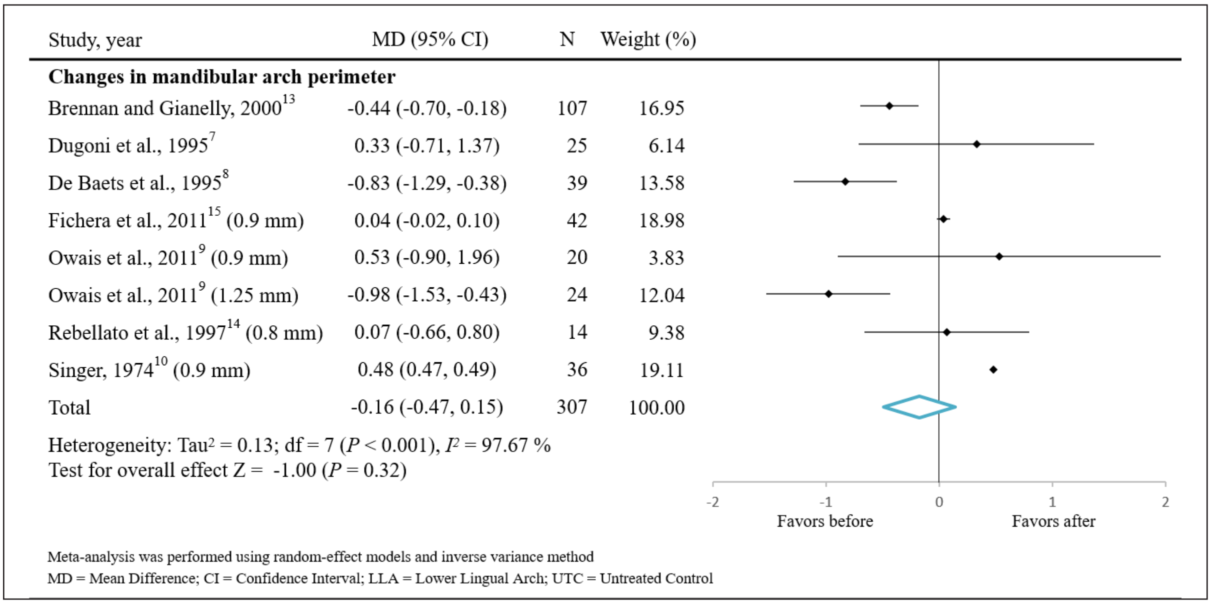

Figure 5. Forest plots for the changes of mandibular arch perimeter comparing pre- and postplacement of lower lingual arch (before and after LLA).

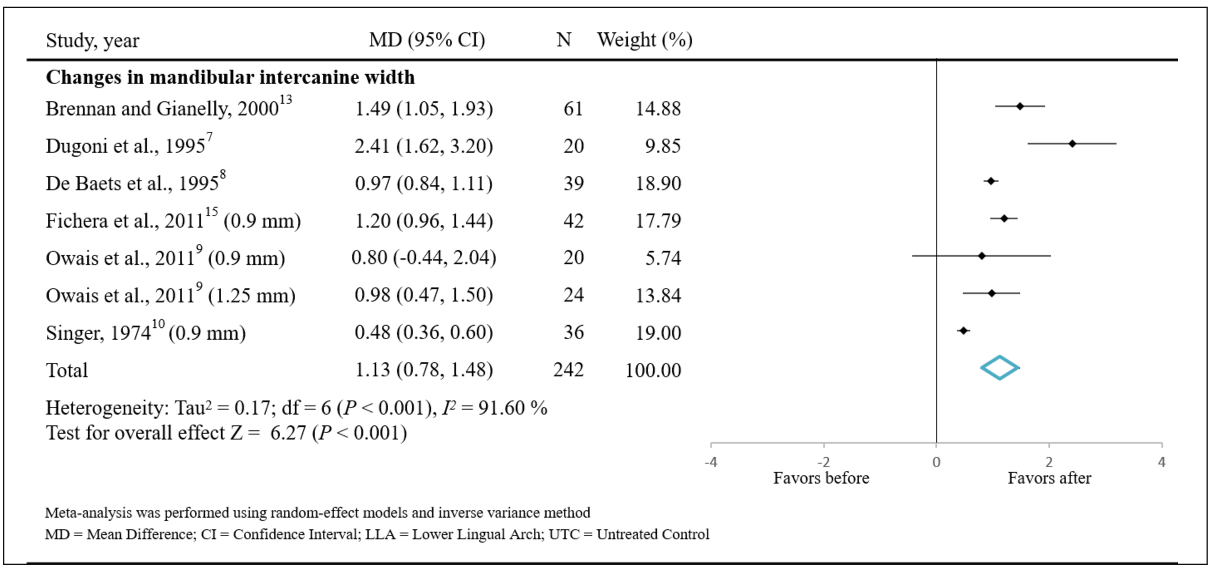

Figure 6. Forest plots for the changes of mandibular intercanine width comparing pre- and postplacement of lower lingual arch (before and after LLA).

fossa. Removing the Rebellato et al.[14] data didn't change the direction of the meta-analysis (MD equals 0.56 mm; 95% CI equals -0.07 to 1.19; *P*=.081). Therefore, the reference point for this outcome did not affect the meta-analysis results on mandibular intermolar width changes between LLA group and UTCs.

Meta-analysis results on mandibular arch dimensional changes comparing pre- and postplacement of LLA: *arch perimeter.* Four retrospective cohort studies, two RCTs, and one nonrandomized clinical trial were included for pre- and post-analysis for the changes of arch perimeter following LLA. The meta-analysis showed that there was a nonsignificant decrease of 0.16 mm in the arch perimeter after LLA (N equals 307; MD equals -0.16 mm; 95% CI equals -0.47 to 0.15; *P*=0.32; Figure 5). The quality of evidence was assessed as very low due to different study designs included and high heterogeneity (I^2 equals 97.67 percent).

Intercanine width. Four retrospective cohort studies, one RCT, and one nonrandomized clinical trial were included for pre- and post-analysis for the changes of intercanine width following LLA. The meta-analysis showed that there was a significant increase of 1.13 mm in intercanine width after LLA (N equals 242; MD equals 1.13 mm; 95% CI equals 0.78 to 1.48; *P*<.001; Figure 6). The quality of evidence was assessed as very low due to different study designs included and high heterogeneity (I^2 equals 91.60 percent).

Intermolar width. Three retrospective cohort studies, two RCTs, and one nonrandomized clinical trial had sufficient data for pre- and post-analysis for the change of intermolar width following LLA. The meta-analysis showed that there was a significant increase in intermolar width after LLA (N equals 268; MD equals 1.08 mm; 95% CI equals 0.39 to 1.77; *P*=.002; Figure 7). The heterogeneity (I^2) was 98.13 percent. In two studies,[7,8] some of the participants received maxillary headgear or two-by-four appliance concomitant with LLA. In order to assess the potential interarch effect reported by Funk,[36] we performed a sensitivity analysis that excluded the two studies.[7,8] The resulting meta-analysis showed an attenuated increase in intermolar width for the LLA participants without maxillary arch treatment (N equals 207; MD equals 0.73 mm; 95% CI equals 0.22 to 1.25; *P*=.005; Figure 8). The quality of evidence was assessed as very low due to various study designs included, and high heterogeneity (I^2 equals 89.65 percent). The meta-analysis of the studies[7,8] with maxillary treatment showed a much larger significant increase of 2.08 mm of intermolar width following LLA (N equals 61; MD equals 2.08 mm; 95% CI equals 2.05 to 2.11; *P*<.001) with a low heterogeneity (I^2 equals 0.00 percent; Figure 8).

Discussion

LLA has been widely used in pediatric dentistry and interceptive orthodontics to preserve the leeway space. However, only a limited number of studies have evaluated the effectiveness of LLA to resolve mandibular incisor crowding and/or its effect on mandibular arch dimensions.[7,9,13-15,20] There has been only one SR concerning LLA published in 2010.[20] This paper studied the effects of LLA on mandibular arch dimensions and restricted the search to RCTs. After their full-text review, there was a lack of well-established control groups and insufficient data in four RCTs, which resulted in only two RCTs being qualified for qualitative analysis in the SR. Although the results showed that LLA was an effective appliance for maintaining arch length and preserving molar anchorage during the transition from primary dentition to permanent dentition, they were not able to provide quantitative results from the meta-analysis.

Results in context with previous studies and implications for practice. The present systematic review has a meta-analysis and provides quantitative results from seven included articles, with a total of 307 participants treated with LLA and 74 untreated participants serving as controls. This review was more complete than the only other systematic review[20] since we included other studies in addition to RCTs. We found that three retrospective cohort studies[7,8,13] reported the resolution of mandibular incisor crowding based on pre- and post-placement of LLA. Although one of these studies used different scales to measure mandibular incisor crowding, the findings were comparable. We found that up to 5.10 mm of mandibular incisor crowding can be resolved by LLA (*P*=.001; Figure 2) from the two studies, and the

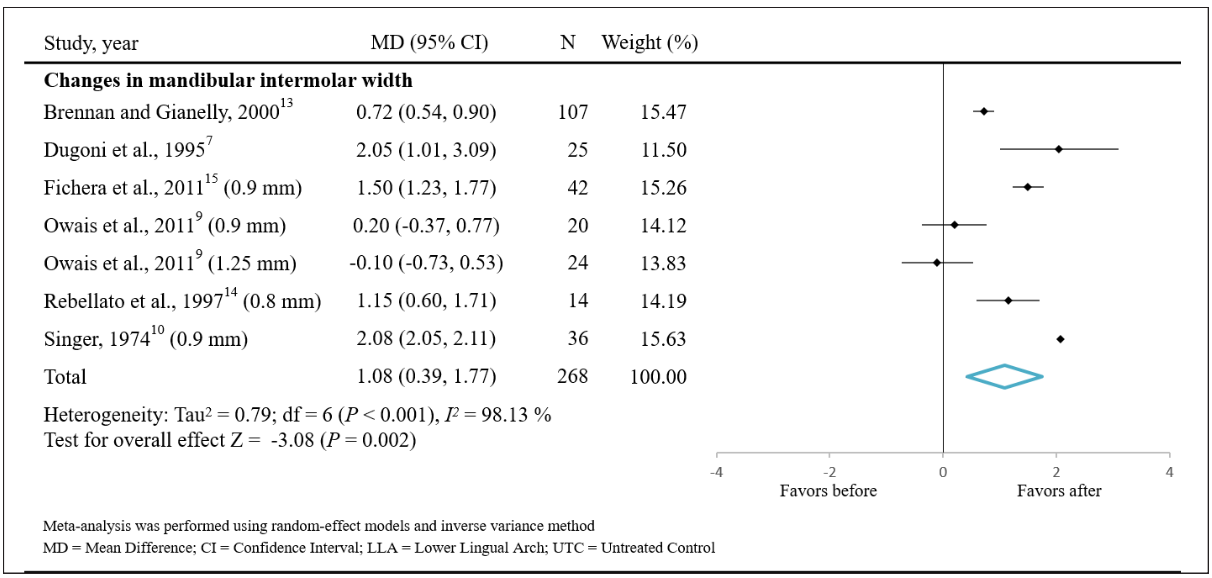

Figure 7. Forest plots for the changes of mandibular intermolar width comparing pre- and postplacement of lower lingual arch (before and after LLA).

third pre- and post-placement LLA study[13] showed a significant resolution of 5.0±2.1 mm of mandibular incisor crowding. This is the first quantitative measure from a systematic review of how much mandibular incisor crowding can be resolved using an LLA.

For the first time, using a meta-analysis, we assessed whether using the LLA compared to UTCs significantly altered mandibular arch perimeter in three studies[9,14,15] or arch length in two studies.[9,14] The meta-analyses showed that there were no significant differences in arch perimeter and arch length between these groups for both outcome measures (P=0.20 and P=0.87, respectively; Figure 3). These statistically nonsignificant changes in mandibular arch perimeter agreed with the data from the seven studies evaluating arch changes from pre- and post-placement of LLA (P=0.32; Figure 5).[7-10,13-15] Most of these seven study's post treatment data was collected after 12 years of age (Table 1). In a normal developing dentition, mandibular arch length decreases up to 3.0 mm between seven and 13 years old due to the loss of leeway space and the mesial drifting of permanent first molars.[37,38] On the other hand, mandibular arch perimeter dimension is mostly established by the age of eight.[9] This is attributed to the eruption of the permanent anterior teeth, followed by a continuous decrease (2.4 mm in males and 3.2 mm in females) due to mesial drifting of permanent first molars from the age of eight to 13.[39] We then can conjecture a reason that LLA is effective in preserving leeway space, maintaining arch perimeter and reducing the amount of space loss from these analyses. Figure 3 showed that LLA-treated patients in the mixed dentition will not differ from UTCs in the amount of mandibular arch perimeter and arch length changes seen over time. Thus, over time, the LLA must prevent most of the mesial drifting of the permanent first molars and maintains the arch perimeter and arch length that normally decreases.

Our findings show a statistically significant increase in intercanine width (Figures 4 and 6), but the arch perimeter and arch length did not increase significantly (Figures 3 and 5). We theorize the LLA was effective in resolving mandibular incisor crowding, as the LLA most likely allows the mandibular incisors to align using some of the primary tooth canine space, which then causes the permanent canines to shift distally into the leeway space during the mixed dentition. The LLA prevents the incisors from tipping or collapsing lingually, as most likely would happen in crowded cases. The distal drifting of the canines will result in a greater intercanine distance that is not a true arch expansion in the canine area. We believe this is the first time a meta-analyses shows nonsignificant arch perimeter and arch length changes yet significant changes in the intercanine width measurements.

The results related to intermolar width are more difficult to explain. We found that there was a significant increase of 0.69 mm in intermolar width in LLA group compared to UTCs[9,14,15] (Figure 4) and a significant increase of 1.08 mm in intermolar width from the pre- and postplacement LLA data,[7,9,10,13-15] (Figure 7) but there was a nonsignificant change in arch perimeter after the placement of LLA (Figures 3 and 5). These findings can possibly be explained by the "interarch effect" theory, as described by Funk.[36] Funk found that utilizing any single arch appliance, like headgear, would cause an interarch effect on the opposite arch. They theorized that the mandibular incisors and molars would move distally with the distal movement of the maxillary molar induced by maxillary headgear.[36] Although Singer[10] reported that headgear alone did not account for all the distal changes in the permanent mandibular molars, they found that LLA may have enhanced the distal movement of incisors and molars along with maxillary headgear.[10] In the seven studies that we included for meta-analysis, only Dugoni et al.[7] and Singer[10] had included

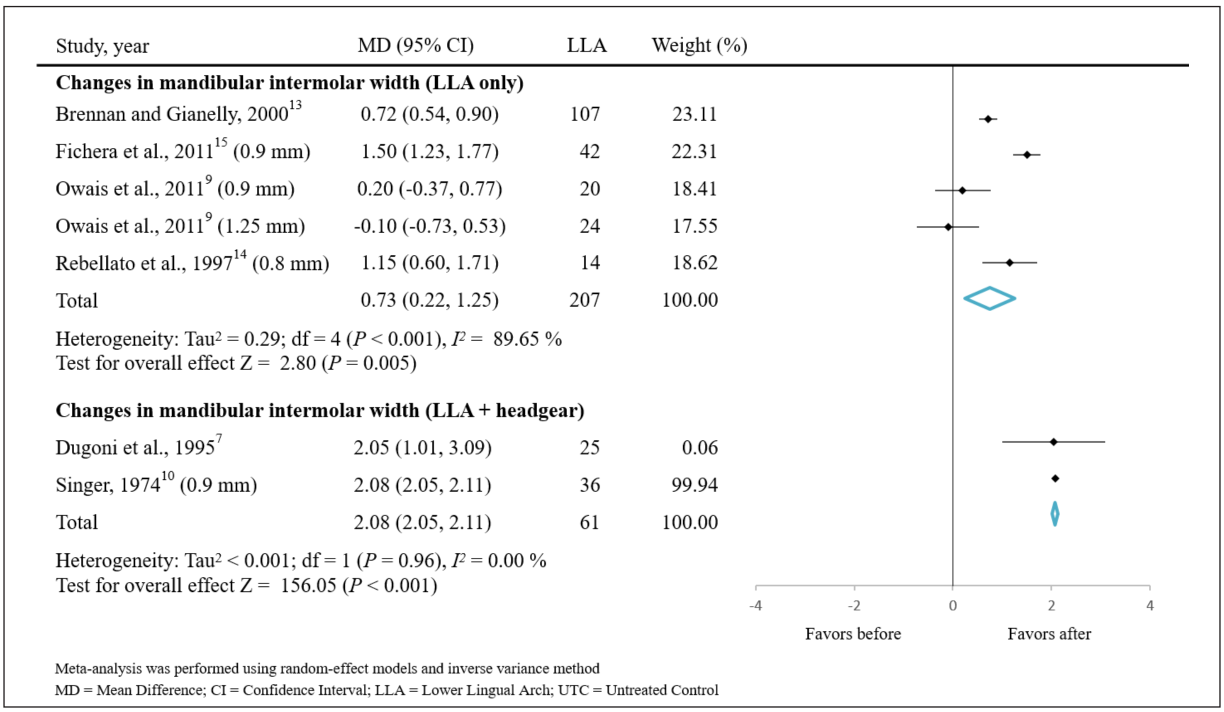

Figure 8. Sensitivity analysis: Forest plots for the changes of mandibular intermolar width comparing pre- and postplacement of lower lingual arch (before and after LLA).

participants that were also treated with headgear. We performed a sensitivity analysis and found that there was a significant increase of 2.08 mm in intermolar width (P<.001) in these two studies,[7,10] whereas there was only an increase of 0.73 mm in intermolar width (P=.005) from the four studies[9,13-15] in which the participants were only treated with LLA (Figure 8). Therefore, we theorized that the LLA combined with headgear may have caused the distal movement of permanent mandibular molars and mandibular incisors; this increased intermolar width significantly, but the arch perimeter remained unchanged due to distal movement of the incisors. This intermolar width expansion was statistically significant but might not be clinically significant in patients who were only treated with LLA, since only 0.73 mm of intermolar movement was found.

Strengths and weaknesses. The current systematic review and meta-analysis followed the guidelines in the Cochrane Handbook for Systematic Reviews of Interventions; additionally, we conducted literature screening, data extraction and synthesis, and assessment of risk of bias using the PRISMA checklist. Two different ROB assessment tools (RoB 2.0 and ROBINS-I) were utilized in this study for RCTs and NRSs. The evaluation of quality of evidence was done using the GRADE approach.

Due to lack of RCTs, the previous SR by Viglianisi[20] was not able to perform quantitative analysis on the effects of LLA on mandibular arch dimension. Thus, the greatest strength of this study was that we didn't restrict the study design to RCTs only. Although meta-analyses that included NRSs present particular challenges because of inherent biases and difference in study designs,[40] we were able to quantify the amount of resolution of mandibular incisor crowding from retrospective cohort studies.

The greatest weakness of this study was lack of RCTs with high quality. In addition, due to the various study designs and high heterogeneity between studies, the quality of evidence was ranked from low to very low. Restricted to the variance of study design and the studied patient pools, the forest plots of the assessed outcome showed high inconsistency and low precision. For example, in the forest plots on mandibular arch perimeter and arch length (Figure 3), although each individual arm showed statistical significance, the high inconsistency/heterogeneity between studies resulted in nonsignificant results. Another weakness was that we have limited response from authors[9,10] who were contacted to clarify the data.

Implications for future research. It would be helpful if authors of future research can follow a consistent way of reporting their data on arch dimensional changes. We had to recompute some data in one article[9] to get the true changes in arch width. In addition, the inconsistencies and lack of standardization on the definitions of arch dimensions made it difficult to extract the correct data. Some defined arch length[9,14,15] or incisor-to-molar distance[10] and really were referring to arch perimeter. Some studies used the term "arch depth" in place of "arch length," which makes data extraction difficult.[9,14] A standard definition of terms is recommended for future research. Furthermore, well-constructed RCTs adopting a standardized approach, such as the CONSORT checklist, are needed in the future to verify the results from what we concluded via the meta-analyses of NRSs.

Conclusions

Based on this study's results, the following conclusions can be made:

1. This systematic review showed that there is very limited low-quality evidence available to support the effect of lower lingual arch use in resolving mild to moderate incisor crowding.
2. The average resolution of mandibular incisor crowding after placement of LLA was 5.10 mm, but the quality of evidence was assessed as very low.
3. Dimensional changes in arch perimeter and arch length were not significantly different between LLA and control groups over the experimental/observational period.
4. There were significant increases in intercanine and intermolar width in LLA group compared to that in UTCs, and there were significant increases in intercanine and intermolar width from data in the pre- and postplacement of LLA.

Acknowledgment

The authors wish to thank Mary Ann Williams, MSLS, at Health Sciences and Human Services Library for her assistance in establishment of literature searching strategy.

References

1. Dean JA. Managing the developing occlusion. In: McDonald RE, Avery DR, eds. Dentistry for the Child and Adolescent. 9th ed. Philadelphia, Pa., USA: CV Mosby Co; 2016:417.
2. Gianelly AA. Leeway space and the resolution of crowding in the mixed dentition. Semin Orthod 1995;1(3): 188-94.
3. Nance HN. The limitations of orthodontic treatment; mixed dentition diagnosis and treatment. Am J Orthod 1947;33(4):177-223.
4. Allen TR, Trojan TM, Harris EF. Evidence favoring a secular reduction in mandibular leeway space. Angle Orthod 2017;87(4):576-82.
5. Proffit WR. Early stage of development. In: Proffit WR, Fields HW, Sarver DM, eds. Contemporary Orthodontics. 2nd ed. St. Louis, Mo., USA: Mosby Year Book; 1993:83.
6. Hahn GW, Kloehn SJ, Lasher MC, et al. Treatment of malocclusion in the mixed dentition: a panel discussion. Angle Orthod 1950;20(2):74-115.
7. Dugoni SA, Lee JS, Varela J, Dugoni AA. Early mixed dentition treatment: postretention evaluation of stability and relapse. Angle Orthod 1995;65(5):311-20.
8. De Baets J, Chiarini M. The pseudo-Class I: a newly defined type of malocclusion. J Clin Orthod 1995;29(2): 73-88.
9. Owais AI, Rousan ME, Badran SA, Abu Alhaija ES. Effectiveness of a lower lingual arch as a space holding device. Eur J Orthod 2011;33(1):37-42.
10. Singer J. The effect of the passive lingual archwire on the lower denture. Angle Orthod 1974;44(2):146-55.
11. Hudson AP, Harris AM, Mohamed N, Joubert J. Use of the passive lower lingual arch in the management of anterior mandibular crowding in the mixed dentition. SADJ 2013;68(3):16-9, 114.
12. Bell RA, Sonis A. Space supervision and guidance of eruption in management of lower transitional crowding: a non-extraction approach. Semin Orthod 2014;20(1): 16-35.

13. Brennan MM, Gianelly AA. The use of the lingual arch in the mixed dentition to resolve incisor crowding. Am J Orthod Dentofacial Orthop 2000;117(1):81-5.
14. Rebellato J, Lindauer SJ, Rubenstein LK, et al. Lower arch perimeter preservation using the lingual arch. Am J Orthod Dentofacial Orthop 1997;112(4):449-56.
15. Fichera G, Greco M, Leonardi R. Effectiveness of the passive lingual arch for E space maintenance in subjects with anterior or posterior rotation of the mandible: a retrospective study. Med Princ Pract 2011;20(2):165-70.
16. Vyas MB, Hantodkar N. Resolving mandibular arch discrepancy through utilization of leeway space. Contemp Clin Dent 2011;2(2):115-8.
17. Senior W. A lingual arch for intruding and uprighting lower incisors. J Clin Orthod 2003;37(6):302-6.
18. Mills CM, McCulloch KJ. Case report: modified use of the Jasper Jumper appliance in a skeletal Class II mixed dentition case requiring palatal expansion. Angle Orthod 1997;67(4):277-82.
19. Celli D, Manente A, DeCarlo A, Deli R. Long-term stability of anterior open bite correction in mixed dentition with a new treatment protocol. Eur J Paediatr Dent 2014;15(2):158-62.
20. Viglianisi A. Effects of lingual arch used as space maintainer on mandibular arch dimension: a systematic review. Am J Orthod Dentofacial Orthop 2010;138(4):382.e1-4; discussion 82-3.
21. Little RM. The irregularity index: a quantitative score of mandibular anterior alignment. Am J Orthod 1975;68(5):554-63.
22. Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. London, UK: The Cochrane Collaboration; 2011.
23. Higgins JP, Sterne JA, Savović J, et al. A revised tool for assessing risk of bias in randomized trials. In: Chandler J, McKenzie J, Boutron I, Welch V, eds. Cochrane Methods. London, UK: The Cochrane Collaboration; 2016.
24. Sterne JA, Hernan MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355:i4919.
25. Kinzinger G, Fritz U, Diedrich P. Combined therapy with pendulum and lingual arch appliances in the early mixed dentition. J Orofac Orthop 2003;64(3):201-13.
26. Little RM, Riedel RA, Stein A. Mandibular arch length increase during the mixed dentition: postretention evaluation of stability and relapse. Am J Orthod Dentofacial Orthop 1990;97(5):393-404.
27. Miotti F. The passive lingual arch in first bicuspid extraction. Angle Orthod 1984;54(2):163-75.
28. Dincer M, Haydar S, Unsal B, Turk T. Space maintainer effects on intercanine arch width and length. J Clin Pediatr Dent 1996;21(1):47-50.
29. Odom WM. Mixed dentition treatment with cervical traction and lower lingual arch. Angle Orthod 1983;53(4):329-42.
30. Sonis A, Ackerman M. E-space preservation: is there a relationship to mandibular second molar impaction? Angle Orthod 2011;81(6):1045-9.
31. Rubin RL, Baccetti T, McNamara JA Jr. Mandibular second molar eruption difficulties related to the maintenance of arch perimeter in the mixed dentition. Am J Orthod Dentofacial Orthop 2012;141(2):146-52.
32. Curtis JW Jr., Camak JB Jr. Complications arising during ingual arch space maintenance. Clin Prev Dent 1983;5(5):22-3.
33. Villalobos FJ, Sinha PK, Nanda RS. Longitudinal assessment of vertical and sagittal control in the mandibular arch by the mandibular fixed lingual arch. Am J Orthod Dentofacial Orthop 2000;118(4):366-70.
34. Letti HC, Rizzatto SM, de Menezes LM, et al. Sagittal changes in lower incisors by the use of lingual arch. Dent Press J Orthod 2013;18(3):29-34.
35. Cuijpers P, Weitz E, Cristea IA, Twisk J. Pre-post effect sizes should be avoided in meta-analyses. Epidemiol Psychiatr Sci 2017;26(4):364-68.
36. Funk AC. Mandibular response to headgear therapy and its clinical significance. Am J Orthod 1967;53(3):182-216.
37. Moorrees CF, Gron AM, Lebret LM, Yen PK, Frohlich FJ. Growth studies of the dentition: a review. Am J Orthod 1969;55(6):600-16.
38. Thilander B. Dentoalveolar development in subjects with normal occlusion: a longitudinal study between the ages of 5 and 31 years. Eur J Orthod 2009;31(2):109-20.
39. Bishara SE, Jakobsen JR, Treder J, Nowak A. Arch length changes from 6 weeks to 45 years. Angle Orthod 1998;68(1):69-74.
40. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283(15):2008-12.

Electronic Appendix

Section 1

Search strategy. PubMed searched strategy last search March 2018.

1. (Orthodontics [MeSH Terms])
2. (lingual [All Fields] OR Nance [All Fields])
3. (arch [All Fields] OR arches [All Fields])

#1 AND #2 AND #3

Section 2

Quality of evidence evaluated using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system.

Pre-LLA compared to Post-LLA in resolving mandibular incisor crowding
Bibliography:

Certainty assessment							Summary of findings				
№ of participants (studies) Follow-up	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Publication bias**	**Overall certainty of evidence**	**Study event rates (%)**		**Relative effect (95% CI)**	**Anticipated absolute effects**	
							With Post-LLA	**With Pre-LLA**		**Risk with Post-LLA**	**Risk difference with Pre-LLA**
Manidublar incisor crowding (follow up: mean 5 years)											
59 (2 observational studies)	serious [a]	very serious [b]	not serious	not serious	all plausible residual confounding would reduce the demonstrated effect	⊕◯◯◯ VERY LOW	0	59	-		MD **5.1 mm higher** (2.15 higher to 8.05 higher)

CI: Confidence interval; **MD:** Mean difference

Explanations

a. The ROBs were ranked from moderate to serious using ROBINS-I tools.
b. The heterogeneity was statistical-significantly high.

Pre-LLA compared to Post-LLA for Mandibular arch dimensions

Bibliography:

Certainty assessment							Summary of findings				
№ of participants (studies) Follow-up	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Publication bias**	**Overall certainty of evidence**	**Study event rates (%)**		**Relative effect (95% CI)**	**Anticipated absolute effects**	
							With Post-LLA	**With Pre-LLA**		**Risk with Post-LLA**	**Risk difference with Pre-LLA**
Arch perimeter (ffollow up: mean 5 years)											
307 (7 observational studies)	serious [a]	very serious [b]	not serious	not serious	all plausible residual confounding would reduce the demonstrated effect	⊕○○○ VERY LOW		307	-		MD **0.16 lower** (0.47 lower to 0.15 higher)
Intercanine widitth (follow up: mean 5 years)											
242 (6 observational studies)	serious [a]	very serious [b]	not serious	not serious	all plausible residual confounding would reduce the demonstrated effect	⊕○○○ VERY LOW		242	-		MD **1.13 mm higher** (0.78 higher to 1.48 higher)
Intermolar width (follow up: mean 5 years)											
268 (6 observational studies)	serious [a]	very serious [b]	not serious	not serious	all plausible residual confounding would reduce the demonstrated effect	⊕○○○ VERY LOW		268	-		MD **1.08 mm higher** (0.39 higher to 1.77 higher)

CI: Confidence interval; **MD:** Mean difference

Explanations

a. Multiple RCTs and NRSs were ranked moderate to serious in ROB.
b. The heterogeneity was statistically-significantly high.

LLA compared to Control for Mandibular arch dimensions

Bibliography:

Certainty assessment							Summary of findings				
№ of participants (studies) Follow-up	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Publication bias**	**Overall certainty of evidence**	**Study event rates (%)**		**Relative effect (95% CI)**	**Anticipated absolute effects**	
							With Control	**With LLA**		**Risk with Control**	**Risk difference with LLA**
Arch perimeter (follow up: mean 5 years)											
180 (3 observational studies)	serious [a]	very serious [b]	not serious	not serious	all plausible residual confounding would reduce the demonstrated effect	⊕◯◯◯ VERY LOW	80	100	-		MD **0.97 mm higher** (0.5 lower to 2.44 higher)
Arch length (follow up: mean 5 years)											
120 (2 RCTs)	serious [c]	very serious [b]	not serious	not serious	all plausible residual confounding would reduce the demonstrated effect	⊕⊕◯◯ LOW	62	58	-		MD **0.09 mm higher** (0.93 lower to 1.1 higher)
Intercanine width (follow up: mean 5 years)											
150 (2 observational studies)	serious [d]	not serious	not serious	not serious	all plausible residual confounding would reduce the demonstrated effect	⊕⊕◯◯ LOW	64	86	-		MD **0.79 mm higher** (0.44 higher to 1.14 higher)
Intermolar width (follow up: mean 5 years)											
180 (3 observational studies)	serious [a]	very serious [b]	not serious	not serious	all plausible residual confounding would reduce the demonstrated effect	⊕◯◯◯ VERY LOW	80	100	-		MD **0.69 mm higher** (0.23 higher to 1.16 higher)

CI: Confidence interval; **MD:** Mean difference

Explanations

a. Two RCTs had Some Concerns in ROB and one NRS was ranked serious in ROB.
b. The heterogeneity was statistically-significantly high.
c. Two RCTs had Some Concerns in ROB.
d. One RCT had Some Concerns in ROB and One NRS was ranked serious in ROB.