Int. J. Oral Maxillofac. Surg. 2019; 48: 364–372
https://doi.org/10.1016/j.ijom.2018.07.014, available online at https://www.sciencedirect.com

International Journal of Oral & Maxillofacial Surgery

Systematic Review
Pre-implant Surgery

Quality assessment of systematic reviews on vertical bone regeneration

J. M. Saletta[1], J. J. Garcia[1,2], J. M. M. Caramês[3,4,5], H. Schliephake[6], D. N. da Silva Marques[4,5,7]

[1]Implant Department, Universidad Europea de Madrid, Madrid, Spain; [2]CIRO, Madrid, Spain; [3]Oral Surgery and Implant Department, Faculdade de Medicina Dentária, Universidade de Lisboa, Lisbon, Portugal; [4]Implantology Institute, Lisbon, Portugal; [5]LIBPhys-FCT UID/FIS/04559/2013, Faculdade de Medicina Dentária, Universidade de Lisboa, Lisbon, Portugal; [6]Department of Oral and Maxillofacial Surgery, University Medicine, George-Augusta-University, Göttingen, Germany; [7]Centro de Estudos de Medicina Dentária Baseada na Evidência, Faculdade de Medicina Dentária, Universidade de Lisboa, Lisbon, Portugal

J. M. Saletta, J. J. Garcia, J. M. M. Caramês, H. Schliephake, D. N. da Silva Marques[4]: Quality assessment of systematic reviews on vertical bone regeneration. Int. J. Oral Maxillofac. Surg. 2019; 48: 364–372.

Abstract. The aim of this study was to evaluate and compare the quality of systematic reviews of vertical bone regeneration techniques, using two quality-assessment tools (AMSTAR and ROBIS). An electronic literature search was conducted to identify systematic reviews or meta-analyses that would evaluate at least one of the following outcomes: implant survival, success rates, complications or bone gain after vertical ridge augmentation. Methodological quality assessment was performed by two independent evaluators. Results were compared between reviewers, and reliability measures were calculated using the Holsti's method® and Cohen's kappa. Seventeen systematic reviews were included, of which seven presented meta-analysis. Mean ±95% confidence interval AMSTAR score was 6.35 [4.74;7.97], with higher scores being correlated with a smaller risk of bias (Pearson's correlation coefficient = −0.84; $P < 0.01$). Cohen's inter-examiner kappa showed substantial agreement for both checklists. From the available evidence, we ascertained that, regardless of the technique used, it is possible to obtain vertical bone gains. Implant success in regenerated areas was similar to implants placed in pristine bone with results equating between 61.5% and 100% with guided bone regeneration being considered the most predictable technique regarding bone stability, while distraction osteogenesis achieved the biggest bone gains with the highest risk of possible complications.

Key words: vertical bone augmentation; distraction osteogenesis; bone regeneration; block graft.

Accepted for publication
Available online 21 August 2018

It is estimated that over 40,000 research articles in the field of dentistry are published each year, and with the increasing number of publications it is not feasible for the clinician in his or her daily practice to evaluate the quality of all the primary evidence published on the subject[1]. Consequently, in recent years the number of systematic reviews (SRs) in the field of dentistry has dramatically increased[2]; over 1188 SRs on oral health issues were published between 1991 and 2012[3], equating to more than 50 per year[3,4]. This secondary evidence follows strict methodology in order to reduce systematic and random errors, thus allowing for the production of a more comprehensive and objective view of the oral health research literature[1].

The rationale for SR literature is grounded firmly on several premises. The first is to reduce and refine large quantities of information into manageable data. Through critical appraisal

and synthesis, an SR separates potentially biased articles from those with sound evidence via quality-assessment tools[5–8], thus establishing a useful decision-making mechanism for clinicians[1]. Secondly, decision makers can and should integrate the critical secondary information available while estimating the variables and outcomes included in their economic and biomedical analysis[7]. Thirdly, an SR is more cost-effective and time efficient than embarking on a new clinical study, provided there is enough evidence available to support a clinical treatment strategy[7].

With the increasing number of published SRs, several tools have been proposed, either as guidelines to follow[8] while writing, appraising the quality of data included[9,10], or even the quality of the SR itself[11,12], in which possible methodological errors are highlighted.

Vertical alveolar bone loss in partially edentulous patients presents several limitations imposed by the remaining anatomical structures and technical restrictions regarding crown–implant ratio that could affect the implant success rate thus increasing the difficulty of implant treatment[13]. Several alveolar reconstruction techniques have been proposed in order to allow simultaneous or deferred implant placement. These techniques include: guided bone regeneration (GBR); bone block grafts (IBG or OBG); distraction osteogenesis (DO); or a combination of the above. Several materials may be used in these procedures including autografts, allografts, xenografts and alloplasts, as well as different barrier membranes or osteosynthesis materials[14,15].

The aim of this study was to evaluate the evidence available for vertical bone regeneration techniques, while determining variability between assessors using different quality assessment tools. In detail, the objectives are as follows: (1) to appraise the quality of the evidence published regarding vertical augmentation techniques prior to or simultaneous with implant placement in patients, using methodological quality assessment tools for systematic reviews (AMSTAR)[11] and risk of bias in systematic reviews (ROBIS)[12] tools; (2) to evaluate the inter-rater agreement between two different evaluators, thus assessing the reliability while using quality-appraisal tools; (3) to synthesize the evidence of the data extracted from the SRs, taking into account the quality of the appraisal, journal impact factor and date of publication regarding the different techniques described in the articles included.

MATERIAL AND METHODS

Search and screening process

A literature search of several electronic databases, including MEDLINE, EMBASE, the Cochrane Central Register of Controlled Trials and the Cochrane Oral Health Group Trials Register, for SRs or meta-analyses written between January 2006 and March 2017 was conducted by two independent reviewers (J.S. and D. M.).

For the PubMed library combinations of controlled terms (MeSH and EMTREE) and keywords were used whenever possible. Search terms were used; where ‘‘[mh]’’ represented the MeSH terms and ‘‘[tiab]’’ represented the title and/or abstract. In addition; other terms not indexed as MeSH and filters were applied. As such; the key terms used were as follows:

PubMed library

- DO: (distraction osteogenesis [MeSH terms]) OR distraction osteogenesis) OR osteogenesis, distraction [MeSH terms]) AND dimension, vertical [MeSH terms]) OR alveolar bone atrophy [MeSH terms]
- GBR: (bone regeneration [MeSH terms]) OR allogenic graft OR xenogenic graft OR synthetic graft, bone replacement [MeSH terms]) AND dimension, vertical [MeSH terms]) OR alveolar bone atrophy [MeSH terms]) OR alveolar bone loss [MeSH terms]
- OBG: (onlay [MeSH terms]) AND bone regeneration [MeSH terms]) AND dimension, vertical [MeSH terms]) OR alveolar bone atrophy [MeSH terms]) OR alveolar bone loss [MeSH terms]
- IBG: (inlay [MeSH terms]) AND bone regeneration [MeSH terms]) AND dimension, vertical [MeSH terms]) OR alveolar bone atrophy [MeSH terms]) OR alveolar bone loss [MeSH terms]

The search strategy was adapted to each database, following the guidelines used for Medline searches. Additionally, a manual search of periodontology- and implantology-related journals, including *Journal of Dental Research*, *Journal of Clinical Periodontology*, *Journal of Periodontology* and *The International Journal of Periodontics & Restorative Dentistry*, from January 2006 to March 2017, was performed to ensure a thorough screening process. Furthermore, references of included articles were screened to check all available articles.

Eligibility criteria

Articles were included in this qualitative assessment if they clearly met the following inclusion criterion: SR or meta-analysis aimed at appraising vertical bone regeneration techniques for implant placement, evaluating at least one of these outcomes: implant survival or success rate, complications or bone gain after vertical ridge augmentation.

Narrative literature reviews were excluded (defined as those without a systematic approach to literature search and selection). References of the articles included were also checked for studies that fulfilled the present inclusion criteria.

Data collection and quality assessment

For each SR included, the following data were extracted: authors, year of publication, research question/objective, number of studies included, outcome measures and journal of publication.

Qualitative evaluation was conducted using different tools, as proposed by recent articles in the periodontology and oral implantology fields[16,17]. First, the checklist advocated by Whiting et al.[12], which consists of 21 questions that assess four different domains (study eligibility, identification and selection, data collection and appraisal, and synthesis and findings), was used, which should enable overall evaluation of the risk of bias. In addition, the AMSTAR tool[11] was used to study the quality of the available evidence on vertical augmentation, thus allowing comparison between both tools. Each article included in this study was appraised independently by two reviewers (D.M. and J. S.), and inter-rater agreement determined.

The AMSTAR checklist features 11 items that can be classified into one of four options: ‘1’ indicating ‘yes’, ‘2’ indicating ‘no’, ‘3’ indicating ‘can’t answer’ and ‘4’ ‘not applicable’. Only option 1 (‘yes’) items generate a score. Therefore, each article could obtain a score between 0 (no criteria) and 11 (all criteria), which coincides with the original proposals[18,19]. There are no guidelines on the scoring using the AMSTAR system. However, it has been proposed that it can be divided into three levels, depending on the overall score obtained: high (8–11), medium (4–7) and low (0–3)[20].

The recently developed ROBIS[12] tool intends to identify concerns regarding the review process and risk of bias across four domains of SRs – study eligibility criteria, identification and selection of studies,

data collection and study appraisal, and synthesis and findings.

Each domain is classified as having high, unclear or low risk of bias, equating to an overall score for the study with the same options for classification.

The outputs of the AMSTAR and ROBIS tools were compared between reviewers, and reliability measures were calculated for all the reviews included by applying Holsti's method® and Cohen's kappa (κ) with SPSS software (IBM, SPSS Statistics for Windows, version 24.0; IBM Corp., Armonk, NY, USA). Cohen's kappa (k) coefficient ± asymptotic standard error (ASE) was used for both checklists, to evaluate the inter-rater agreement for qualitative items[21]. We assessed questions with kappa values of <0 as having less than chance agreement, 0.01–0.20 as slight agreement, 0.21–0.40 as fair agreement, 0.41–0.60 as moderate agreement, 0.61–0.80 as substantial agreement and 0.81–0.99 as almost perfect agreement[21]. Pearson correlation coefficients (ρ) were used to correlate AMSTAR and ROBIS assessments, journal impact factor and year of publication, with $P < 0.05$ considered as significant. The strength of the resulting correlations was described using established criteria[22]. The results were expressed as mean and 95% confidence interval (CI) unless otherwise noted.

In order to produce a narrative synthesis based on the quality appraisal, after the inter-rater reliability calculations, the discrepancies were resolved through discussion.

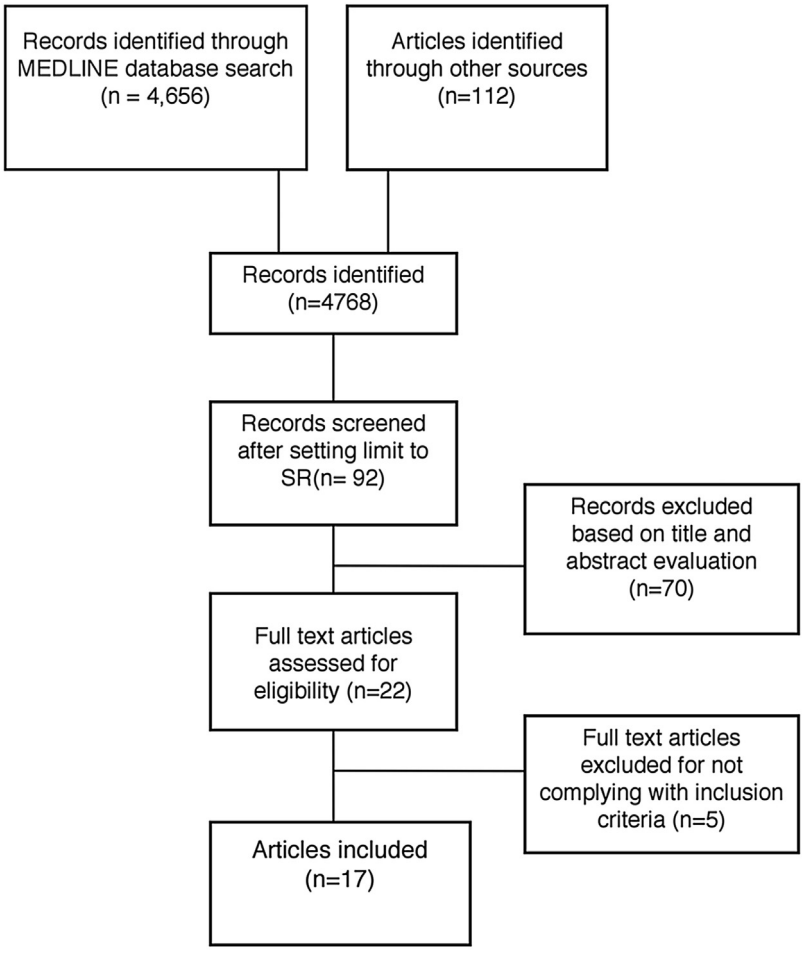

Fig. 1. PRISMA flowchart diagram of the search strategy. The articles are sorted by identification, screening and eligibility.

RESULTS

Literature search

The initial search found 4656 titles in Medline/PubMed, and 112 titles from the other sources. No additional studies were found in the grey literature or in the references of the studies included. Of the initial 92 screened records, 70 were excluded based on title or abstract (56 were not SRs, 11 did not consider vertical augmentation procedures and 3 did not complied with eligibility criteria). After the first evaluation, 22 full articles were selected. After careful reading, five studies were excluded for not complying with the eligibility criteria. Therefore, 17 studies published between 2009 and 2017 were included (Fig. 1).

Study characteristics

Regarding the pre-established outcomes, four articles evaluated the implant failure rate[23–26], defined as implant mobility and removal of stable implants dictated by progressive marginal bone loss or infection,12 evaluated implant survival[26–37], which was presented (when possible) as cumulative survival percentage rate. Any other definitions of implant survival, as described in individual studies, were also considered. Seven described the implant success rate[26,31–33,36,38,39], defined with different criteria between the studies.

Post-surgical complications were evaluated in eight studies[23–26,28,32,35,37], while prosthesis survival was assessed in four[23–25,37].

Bone loss was determined in three studies[32,36,37], while bone gain after vertical augmentation was analysed in 10 studies[23–29,31,32,37].

Of the selected sample, seven of studies presented quantitative results as a meta-analysis[25,30–32,34,36,37], although with large variability in the number of studies included in each review, ranging between 4 and 90.

Four manuscripts only included randomized controlled trials[23–25,37] while the other manuscripts also included controlled clinical trials, as well as prospective and retrospective cohort studies, case–control studies and case reports.

The overall AMSTAR and ROBIS classifications of the studies included are depicted in Table 1. The clinical questions were formulated and organized according to the PICO strategy[40] in 12 of the studies included[23–25,29–32,35–39], the electronic database most frequently used was Medline/PubMed, with 100% referring to such a search; three studies did not present language restrictions[23–25], and eight did not include critical appraisal of the studies included[26–29,34–36,39].

Inter-rater agreement

For the AMSTAR checklist[11], the overall κ coefficient was 0.79 (0.04), and for the ROBIS checklist with the five options[12], it was 0.76 (0.02).

Table 1. Main characteristics and conclusions of the Systematic Reviews selected for the quality assessment.

Authors	Included studies	Journal	AMSTAR	ROBIS	Main results/conclusions
Esposito et al., 2006[23]	13	*Int J Oral Maxillofac Implants*	High	Low	GBR and DO can be used to VRA, but is unclear which is the most efficient
Aghaloo et al., 2007[34]	90	*Int J Oral Maxillofac Implants*	Low	High	ISR: 95.5% GBR; 90.4% OVG; 94.7% DO; 83.8% COG
Esposito et al., 2008[24]	17	*The Cochrane Library*	High	Low	GBR and DO can be used to VRA, but there is insufficient evidence to indicate which is preferable
Rocchietta et al., 2008[26]	26	*J Clin Periodontol*	Low	High	VRA: 2–8 mm; ISR 92.1–100%; ISS: 76.3–97.5%; C: 0–45.5%
Saulacic et al., 2008[27]	20	*Int J Oral Maxillofac Surg*	Low	High	DO: VRA: 6.88 ± 2.52 mm; ISR: 97%
Esposito et al., 2009[25]	13	*The Cochrane Library*	High	Low	GBR and DO can be used to VRA, but there is insufficient evidence to indicate which could be preferable
Waasdorp et al., 2010[28]	9	*Int J Oral Maxillofac Implants*	Medium	High	OBG: VRA: 3–4 mm; ISS: 99.9%; C: 8.5%
Clementini et al., 2011[39]	8	*Oral Implantol (Rome)*	Medium	High	OBG: ISS: 72.8–97%
Klein et al., 2011 (28)	72	*Eur J Oral Implantol*	Medium	High	VRA: 2.0–5.6 mm; ISR: 66.7–100% The screened data did not allow identification of a superior grafting approach
Clementini et al., 2012[38]	8	*Int J Oral Maxillofac Surg*	Medium	High	GBR: ISS: 61.5–100%
Al-Nawas et al., 2014[30]	52/14	*Eur J Oral Implantol*	High	Low	GBR: ISR: 97.4–100%; ISS: 90.3–100%
Milinkovic et al., 2014[35]	68	*Int J Oral Maxillofac Surg*	Medium	High	ISR: GBR 98.9% simultaneous 100% staged; OBG/IBG 96.3%, DO 98.2%; C: GBR 13.1% simultaneous 6.95% staged; OBG/IBG 8.1%; DO 22.4%
Aloy-Prósper et al., 2015[33]	6	*Med Oral Patol Oral Cir Bucal*	Medium	High	Survival and success rates of implants placed in vertically reconstructed ridges with bone block grafts are similar to those placed in native bone, in distracted sites or with GBR
Camps-Font et al, 2016[37]	14	*J Periodontol*	High	Low	Vertical regeneration versus short implants: IF: [OR]: 1.02; 95% [CI] 0.31–3.31; C: [OR]: 8.33; 95% [CI]: 3.85–20.0
Keestra et al., 2016[36]	51	*J Appl Oral Sci*	Medium	High	ISR – mean (range): DO 97.1% (88–100%); IBG 98.5% (95.9–100%); OBG 94.7% (72.8–100%); GBR 99.3% (94.1–100%) Considering studies with 4–5 years follow-up, there seems to be a trend that OBG, DO and GBR are stable for at least 4–5 years
Yun et al, 2016[31]	4	*Int J Oral Maxillofac Implants*	Medium	Unclear	There was not a statistically significant difference between DO and OBG in terms of bone gain and bone resorption
Elnayef et al., 2017[32]	73	*Int J Oral Maxillofac Implants*	High	Low	GBR is the most reliable technique in terms of bone stability (resorption, complication rate and morbidity), while the greatest VRA can be obtained with DO and IBG, although presenting the higher complication rates

C, complications; CI, confidence interval; COG, combinations of onlay, veneer, and interpositional inlay grafting; DO, distraction osteogenesis; GBR, guided bone regeneration; IBG, inlay bone graft; IF, implant failure; ISR, implant survival rate; ISS, implant success rate; OBG, onlay bone graft; OR, odds ratio; OVG, onlay or veneer grafting; VRA, vertical augmentation.

As a secondary measure, the ROBIS answers were aggregated into three options (positive = yes and probably yes; negative = no and probably no; unclear), which corresponded to a κ coefficient of 0.87 (0.02), showing almost perfect inter-examiner agreement[21].

Substantial agreement (>80% or nearly 80%) was observed for the majority of individual items (Table 2) for AMSTAR. However, item 9 (appropriate method to combine scores) presented a lower degree of agreement (70.6%) and kappa parameter (0.46) for all individual results, although the percentage agreement was still considered moderate.

When evaluating the inter-rater agreement, we observed that, for AMSTAR, the overall kappa was 0.79 (0.04), which can be considered as substantial agreement. When using the ROBIS checklist, we observed that aggregating the item answers into three groups (yes, no, can't tell) increased the kappa to 0.87 (0.02), corresponding to an almost perfect agreement in comparison to using five possible answers (yes, probably yes, no, probably no, can't tell) which presented a substantial agreement, with a kappa of 0.76 (0.02). However, when comparing the kappa coefficient for individual items, we observed that the results were similar between the three- and five-group answers for the four domains and for evaluating risk of bias (Table 3), leading to the hypothesis that while aggregating answers into a small number of possible options could lead to higher overall inter-rater agreement, it will not be reflected in the main aim of the ROBIS checklist, which is appraising the quality of the articles in the four available domains and determining the risk of bias.

Table 2. Interobserver agreement per AMSTAR question.

AMSTAR CHECKLIST		κ value	Asymptotic standard error	% AGREEMENT
1	Was a priori design provided?	0.75	0.17	88%
2	Was there duplicate study selection and data extraction?	1	0	100%
3	Was a comprehensive literature search performed?	0.51	0.23	82%
4	Was the status of publication (ie, grey literature) used as an inclusion criterion?	0.90	0.10	94%
5	Was a list of studies (included and excluded) provided?	0.89	0.10	94%
6	Were the characteristics of the included studies provided?	c	0	94%
7	Was the scientific quality of the included studies assessed and documented?	0.65	0.17	82%
8	Was the scientific quality of the included studies used appropriately in formulating conclusions?	1	0	100%
9	Were the methods used to combine the findings of studies appropriate?	0.46	0.18	71%
10	Was the likelihood of publication bias assessed?	0.79	0.12	88%
11	Was the conflict of interest included?	c	0	82%

c, No statistics were computed because one of the raters' answers were constant.

Table 3. Inter-observer agreement per ROBIS question with 5 and 3 possible answers.

ROBIS 5 options				ROBIS 3 options			
QuestioN	κ value	Asymptotic standard error	% Agreement	Question	κ value	Asymptotic standard error	% Agreement
1.1	0.61	0.16	88%	1.1	−0.06	0.04	88%
1.2	0.50	0.23	82%	1.2	c	0	94%
1.3	0.44	0.18	70%	1.3	0.77	0.22	94%
1.4	0.57	0.16	76%	1.4	0.75	0.16	88%
1.5	0.60	0.14	70%	1.5	0.65	0.17	82%
2.1	0.55	0.20	76%	2.1	1	0	100%
2.2	0.69	0.13	76%	2.2	1	0	100%
2.3	0.91	0.09	94%	2.3	1	0	100%
2.4	0.46	0.11	64%	2.4	0.61	0.19	82%
2.5	0.64	0.15	76%	2.5	1	0	100%
3.1	0.77	0.14	88%	3.1	0.82	0.17	94%
3.2	0.90	0.09	94%	3.2	0.78	0.22	94%
3.3	0.71	0.15	82%	3.3	1	0	100%
3.4	0.55	0.15	70%	3.4	0.77	0.14	88%
3.5	0.82	0.12	88%	3.5	0.88	0.12	94%
4.1	0.47	0.18	76%	4.1	0.31	0.15	88%
4.2	0.58	0.13	70%	4.2	0.81	0.12	88%
4.3	1	0	100%	4.3	1	0	100%
4.4	0.66	0.14	76%	4.4	0.69	0.16	82%
4.5	0.79	0.12	88%	4.5	0.88	0.12	94%
4.6	0.91	0.09	94%	4.6	1	0	100%
A	0.60	0.14	70%	A	0.88	0.12	94%
B	0.36	0.16	58%	B	0.72	0.19	88%
C	0.39	0.19	70%	C	0.77	0.22	94%
Domain 1	0.73	0.13	82%	Domain 1	0.73	0.13	82%
Domain 2	0.68	0.15	82%	Domain 2	0.68	0.15	82%
Domain 3	0.88	0.12	94%	Domain 3	0.88	0.12	94%
Domain 4	0.89	0.10	94%	Domain 4	0.89	0.10	94%
Risk of bias	0.89	0.1	94%	Risk of bias	0.89	0.1	94%

c, No statistics were computed because one of the raters' answers were constant.

Quality assessment

After initial appraisal by the two examiners, all areas of disagreement were discussed until a consensus was reached and the final quality assessment scores for AMSTAR obtained.

The mean overall AMSTAR score was 6.35 [4.74; 7.97] out of a maximum score of 11, with 35% of the SRs included achieving a high score (overall score >7). A high negative correlation was found between AMSTAR score and qualitative assessment or risk of bias classification by ROBIS, higher scores (Pearson's correlation coefficient = −0.84 and $P < 0.01$) or higher quality (Pearson's correlation coefficient = −0.86 and $P < 0.01$) being correlated with smaller risk of bias.

The present study also assessed whether the year of publication or the journal citation impact factor correlated with the AMSTAR score. It was found that only the citation impact factor showed a moderate correlation with the AMSTAR score (Pearson's correlation coefficient = 0.54 and $P < 0.05$), a higher impact factor corresponding to a higher AMSTAR score.

Based on individual analysis of the AMSTAR checklist[11], the lowest individual question scores were obtained for the inclusion of grey literature or declaring a conflict of interest according to AMSTAR guidelines (Fig. 2).

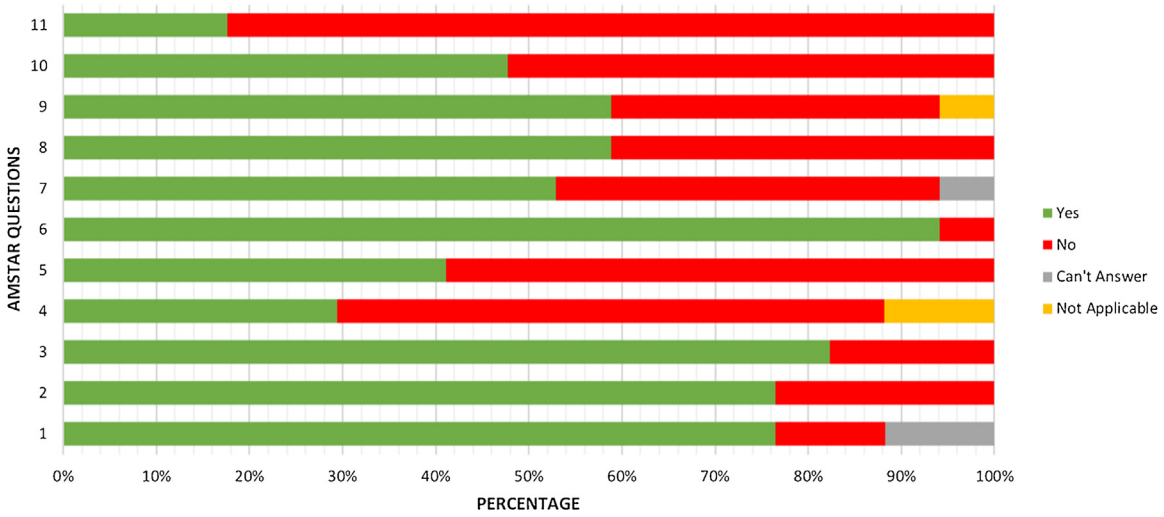

Fig. 2. Overall percentage of positive answers for the individual items of the AMSTAR checklist.

Outcome measures

There is clinical and histological data supporting vertical augmentation techniques for implant placement[26].

Of the SRs analysed, we ascertained that, regardless of the technique used, it is possible to obtain vertical bone gains with values ranging from 2 to 8 mm, depending on the SR[25–29,31,32,35,37], although the SR with the highest number of articles included regarding bone gain, and higher AMSTAR and ROBIS scores, reported bone gain of 4.49mm (95% CI: 3.85–5.14 mm)[32].

When evaluating implant survival and success rates after vertical augmentation procedures, the values ranged from 83.8% to 100%[26–39] and from 61.5% to 100%[26,31–33,36,38,39], respectively.

When evaluating different techniques to vertically augment bone, none of the SRs included was able to clarify which was most efficient[23–26,29,31,32,37] with respect to bone gain or bone resorption.

When assessing post-surgical complications, 47% of the SRs included evaluated this outcome; they referred to a large range of complication rates (0–60%), mostly due to sensory disorders, wound opening, membrane exposure and prosthetic failure. When evaluating the complication rate by technique, it was possible to observe that the DO technique involved the majority of complications, such as lingual vector inclination and loosening of the distractor[32].

DISCUSSION

This article intended to evaluate the methodological quality and risk of bias of SRs published on vertical augmentation techniques prior to implant placement. For that purpose, two different tools, AMSTAR and ROBIS, were used[11,12]. The inter-rater agreement between an experienced evaluator and a clinician was assessed using both scales, and based on the quality assessments, synthesis of the evidence available was produced.

SRs with or without meta-analyses should be regarded as having the highest quality in evidence-based science. Thus, they are fundamental to decision-making in the scientific community, allowing clinicians to perform evidence-based decisions in their clinical practice in a less time-consuming way[40].

In addition, quantitative analysis (meta-analysis) can pool the final results of various studies obtaining an overall estimate while analysing the heterogeneity, and consequently increasing the precision and statistical power of the review[41].

With an increasing number of secondary studies published in the scientific literature presenting contradictory results due to methodological flaws, quality assessment of such reviews should be performed in order to ascertain the quality of the evidence produced[16,17].

Several SRs have been published regarding vertical bone regeneration, although their methodological quality and risk of bias, all important factors in decision-making have never been assessed.

For this purpose, in the present study the authors used two different tools (AMSTAR and ROBIS) in order to evaluate whether there is correlation between the scores obtained in each tool and whether the existence of more than one evaluator can influence the results obtained. The rationale was to assess which is the most efficient and reliable tool for an active clinician to use in daily practice to answer a clinical question.

Based on the obtained results, both tools are reliable for appraising the quality of the SRs included which is in accordance with similar studies[42,43]. In addition, the existence of two different evaluators did not influence the results, as there was high inter-rater agreement ($\kappa = 0.79$ and 0.76 for AMSTAR and ROBIS, respectively). These results are important and novel since, to our knowledge, it the first time that the inter-rater reliability has been assessed while using these tools. Evaluating the individual items in AMSTAR, the lowest percentage of agreement was for item 9 (methods to combine results) which could be related to the experience factor since one of the evaluators (D.M.) already had former experience in publishing secondary evidence. Although the percentage of agreement between reviewers ranged between 70% and 82%, there was some disagreement regarding what could be considered a comprehensive search, thus decreasing the kappa for item 9. And, from the point of view of the authors, the experience factor is fundamental when evaluating the quality of the statistical models used, particularly while performing meta-analysis, which requires advanced statistical knowledge that an average clinician does not use in his or her daily practice[44].

Addressing the individual items in ROBIS, the options with a lower level of agreement were those that related the conclusions to supporting evidence, although the disagreements were mostly related to answers in between positive (yes/probably yes) and negative (no/prob-

ably no), which agrees with a higher kappa, when the ROBIS options are reduced to three. In the authors' opinion, these small disagreements could be again related to the experience factor, but due to the small sample size of SRs included, no clear conclusion could be drawn regarding inter-rater reliability.

The concept of risk of bias is distinct from methodological quality, in that it assesses systemic flaws or limitations in the design, conduct or analysis of research that distort the findings. Although there is some overlap between risk of bias and methodological quality criteria, the majority of items are distinct. Taking this into consideration, the authors evaluated the correlation between AMSTAR and ROBIS, and although they evaluate different criteria, it was possible to ascertain a negative correlation ($P < 0.01$) between them, demonstrating a correlation between methodological flaws and the risk of bias.

Nonetheless, with the plethora of tools and guidance available, there is still confusion over which are the best criteria and tools to assess quality or reporting for consistent standards across reports[8].

The authors evaluated whether the year of publication and the impact factor of the journal in which the SR was published correlated with overall AMSTAR score. It could be expected that more recent publications would relate to higher AMSTAR scores, since the amount of evidence available increases each year. However, when performing the analysis, a significant positive correlation was only found between journal impact factor and AMSTAR score (Pearson correlation coefficient = 0.54 and $P = 0.027$). These results imply that some journals may have higher standards for evidence-based studies on dentistry, which is in agreement with the findings of previous studies[17,45].

While assessing the overall quality of the SRs included, we ascertained that the vast majority are of medium quality, mainly due to the presence of methodological flaws and lower quality of the evidence included.

In our analysis, few of the reviews fulfilled all related items in the checklists, which could be related to the fact that only four of the SRs included were based exclusively on randomized controlled trials, thus decreasing the overall quality due to the inclusion of lower-quality studies.

At present, and although there are some guidelines for performing an SR, in the oral health field it is possible to observe large variability in the methodological design of the published evidence[10]. In order to avoid or minimize methodological errors and the presence of bias in future reviews, a standardized checklist has been proposed[46]. It includes 16 key items from the Cochrane Handbook for Systematic Reviews of Interventions[10], PRISMA[47] and quality-analysis tools[11]. This checklist, which is objective and easy to interpret, may help researchers in the elaboration of future SRs.

Of the 17 SRs analysed, according to the authors, one followed the QUORUM statement[48], five[23–25,31,36] the Cochrane Handbook for Systematic Reviews of Interventions[10], and three[30,32,33] the PRISMA guidelines[47], while one followed the STROBE statement[49]. The use of these guidelines can increase the methodological transparency of the research and consequently lower the risk of bias and increase the quality of scientific evidence.

Regeneration of vertical alveolar bone deficiencies is one of the most difficult clinical scenarios due to anatomical limitations and difficulties in regeneration techniques. Although short implants have been presented as an alternative to these limitations, in a recently published meta-analysis, caution is suggested when using implants shorter than 8 mm (4–7 mm), because they could present a higher risk of failure compared to standard-length implants[50]. The implant–crown ratio would also interfere with the aesthetic and biomechanical outcomes when using these implants[51].

For these reasons, several techniques, such as GBR, bone block grafts (IBG or OBG) and DO, have been proposed to vertically augment the alveolar process. Based on the results of the evaluated SRs, the use of GBR techniques could be considered as the most predictable in terms of bone stability (smaller bone resorption and number of complications) while according to Elnayef et al., DO is the technique which affords the greatest amount of vertical ridge augmentation (mean: 6.84 mm)[32], although Esposito et al. indicate that it is of little use in the presence of thin ridges[24].

Various SRs suggest that the techniques available are suitable for the augmentation of the edentulous ridge, and there is no technique that is more efficient than others[23,24,31,37], although operator-sensitive and postoperative complications are common, varying between sensory disorders, wound opening, membrane exposure and prosthetic failure[37], and that in some cases they can even determine the failure of the regeneration procedure[23,24].

When assessing implant survival, all the SRs included reported similar results to those previously published for implant placement in pristine bone, with survival rates ranging from 83.8% to 100%[25–38]; however, most detected heterogeneity in study design.

The SRs included in this study presented variable methodological quality, with only 41% being classified as high quality. However, some items were poorly attended to in the majority of the SRs included (no comprehensive research strategy, data based on few trials with a small sample of patients with a short follow-up, and exposed to high risk of bias). For this reason, the critical appraisal tools are of mandatory importance when a clinician wants to implement his daily clinical protocols with a new intervention.

To improve this, we suggest the use of specific guidelines for conducting such types of review, and readers should use critical appraisal tools, since they enable assessment of several aspects that increase the scientific validity and strength of the purposed recommendations.

Within the limitations of this study, it can be concluded that vertical bone regeneration techniques present themselves as efficient regarding bone gain and implant survival although they are operator sensitive. Nevertheless, further studies with longer follow-up periods are needed, to better clarify which technique presents the best results for the previously mentioned clinical outcomes.

Funding

This work was supported by the Evidence Based Dentistry Center, Lisbon University and Implant Dentistry Department, European University of Madrid.

Competing interests

The authors declare that they have no conflicts of interest.

Ethical approval

Not applicable.

Patient consent

Not applicable.

References

1. Glenny AM, Esposito M, Coulthard P, Worthington HV. The assessment of systematic reviews in dentistry. *Eur J Oral Sci* 2003;**111**:85–92.
2. Faggion Jr CM, Huda F, Wasiak J. Use of methodological tools for assessing the quality of studies in periodontology and implant

dentistry: a systematic review. *J Clin Periodontol* 2014;**41**:625–31. http://dx.doi.org/10.1111/jcpe.12251.

3. Saltaji H, Cummings GG, Armijo-Olivo S, Major MP, Amin M, Major PW, Hartling L, Flores-Mir C. A descriptive analysis of oral health systematic reviews published 1991-2012: cross sectional study. *PLoS One* 2013;**8**:e74545. http://dx.doi.org/10.1371/journal.pone.0074545.
4. Major MP, Major PW, Flores-Mir C. Benchmarking of reported search and selection methods of systematic reviews by dental speciality. *Evid Based Dent* 2007;**8**:66–70. http://dx.doi.org/10.1038/sj.ebd.6400504.
5. Evans D. Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *J Clin Nurs* 2003;**12**:77–84.
6. Monje A, Wang HL. Interpretation of study design on marginal bone loss in implant dentistry: evidence-based science versus clinical-based experience. *Int J Oral Maxillofac Implants* 2014;**29**:279.
7. Mulrow CD. Rationale for systematic reviews. *BMJ* 1994;**309**:597–9.
8. Pussegoda K, Turner L, Garritty C, Mayhew A, Skidmore B, Stevens A, Boutron I, Sarkis-Onofre R, Bjerre LM, Hróbjartsson A, Altman DG, Moher D. Identifying approaches for assessing methodological and reporting quality of systematic reviews: a descriptive study. *Syst Rev* 2017;**6**:117. http://dx.doi.org/10.1186/s13643-017-0507-6.
9. Critical Appraisal Skills Programme (2018). CASP Checklist. Available at: https://casp-uk.net/casp-tools-checklists/.
10. Higgins JPT, D. G. Altman, and J. A. C. Sterne. Assessing risk of bias in included studies. Cochrane Handbook for Systematic Reviews of Interventions 2011. Available at: http://handbook-5-1.cochrane.org/.
11. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, Porter AC, Tugwell P, Moher D, Bouter LM. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007;**7**:10. http://dx.doi.org/10.1186/1471-2288-7-10.
12. Whiting P, Savovic J, Higgins JP, Caldwell DM, Reeves BC, Shea B, Davies P, Kleijnen J, Churchill R. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 2016;**69**:225–34. http://dx.doi.org/10.1016/j.jclinepi.2015.06.005. ROBIS group.
13. Mecall RA, Rosenfeld AL. Influence of residual ridge resorption patterns on implant fixture placement and tooth position. 1. *Int J Periodontics Restorative Dent* 1991;**11**:8–23.
14. Chiapasco M, Casentini P, Zaniboni M. Bone augmentation procedures in implant dentistry. *Int J Oral Maxillofac Implants* 2009;**24** (Suppl):237–59.
15. Hammerle CH, Jung RE, Feloutzis AA. A systematic review of the survival of implants in bone sites augmented with barrier membranes (guided bone regeneration) in partially edentulous patients. *J Clin Periodontol* 2002;**29**(Suppl 3):223–32. 226-231 discussion.
16. Moraschini V, Luz D, Velloso G, Barboza EDP. Quality assessment of systematic reviews of the significance of keratinized mucosa on implant health. *Int J Oral Maxillofac Surg* 2017;**46**:774–81. http://dx.doi.org/10.1016/j.ijom.2017.02.1274.
17. Monje A, Pommer B. The concept of platform switching to preserve peri-implant bone level: assessment of methodologic quality of systematic reviews. *Int J Oral Maxillofac Implants* 2015;**30**:1084–92. http://dx.doi.org/10.11607/jomi.4103.
18. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, Henry DA, Boers M. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol* 2009;**62**:1013–20. http://dx.doi.org/10.1016/j.jclinepi.2008.10.009.
19. Riado Minguez D, Kowalski M, Vallve Odena M, Longin Pontzen D, Jelicic Kadic A, Jeric M, Dosenovic S, Jakus D, Vrdoljak M, Poklepovic Pericic T, Sapunar D, Puljak L. Methodological and reporting quality of systematic reviews published in the highest ranking journals in the field of pain. *Anesth Analg* 2017. http://dx.doi.org/10.1213/ANE.0000000000002227.
20. Sharif MO, Janjua-Sharif FN, Ali H, Ahmed F. Systematic reviews explained: AMSTAR-how to tell the good from the bad and the ugly. *Oral Health Dent Manag* 2013;**12**:9–16.
21. Landis J, Koch G. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74.
22. Hinkle DE, Wiersma W, Jurs SG. *Applied Statistics for the Behavioral Sciences*. 5th wd. Boston: Houghton: Mifflin; 2003.
23. Esposito M, Grusovin MG, Coulthard P, Worthington HV. The efficacy of various bone augmentation procedures for dental implants: a Cochrane systematic review of randomized controlled clinical trials. *Int J Oral Maxillofac Implants* 2006;**21**:696–710.
24. Esposito M, Grusovin MG, Kwan S, Worthington HV, Coulthard P. Interventions for replacing missing teeth: bone augmentation techniques for dental implant treatment. *Cochrane Database Syst Rev* 2008: CD003607. http://dx.doi.org/10.1002/14651858.CD003607.pub3.
25. Esposito M, Grusovin MG, Felice P, Karatzopoulos G, Worthington HV, Coulthard P. Interventions for replacing missing teeth: horizontal and vertical bone augmentation techniques for dental implant treatment. *Cochrane Database Syst Rev* 2009: CD003607. http://dx.doi.org/10.1002/14651858.CD003607.pub4.
26. Rocchietta I, Fontana F, Simion M. Clinical outcomes of vertical bone augmentation to enable dental implant placement: a systematic review. *J Clin Periodontol* 2008;**35**:203–15. http://dx.doi.org/10.1111/j.1600-051X.2008.01271.x.
27. Saulacic N, Iizuka T, Martin MS, Garcia AG. Alveolar distraction osteogenesis: a systematic review. *Int J Oral Maxillofac Surg* 2008;**37**:1–7. http://dx.doi.org/10.1016/j.ijom.2007.07.020.
28. Waasdorp J, Reynolds MA. Allogeneic bone onlay grafts for alveolar ridge augmentation: a systematic review. *Int J Oral Maxillofac Implants* 2010;**25**:525–31.
29. Oliver Klein M, Al-Nawas Bilal. For which clinical indications in dental implantology is the use of bone substitute materials scientifically substantiated? Systematic review, consensus statements and recommendations of the 1st DGI Consensus Conference in September 2010. *Eur J Oral Implantol* 2011;**4**.
30. Al-Nawas B, Schiegnitz E. Augmentation procedures using bone substitute materials or autogenous bone – a systematic review and meta-analysis. *Eur J Oral Implantol* 2014;**7**(Suppl 2):S219–234.
31. Yun KI, Choi H, Wright RF, Ahn HS, Chang BM, Kim HJ. Efficacy of alveolar vertical distraction osteogenesis and autogenous bone grafting for dental implants: systematic review and meta-analysis. *Int J Oral Maxillofac Implants* 2016;**31**:26–36. http://dx.doi.org/10.11607/jomi.4479.
32. Elnayef B, Monje A, Gargallo-Albiol J, Galindo-Moreno P, Wang HL, Hernandez-Alfaro F. Vertical ridge augmentation in the atrophic mandible: a systematic review and meta-analysis. *Int J Oral Maxillofac Implants* 2017;**32**:291–312. http://dx.doi.org/10.11607/jomi.4861.
33. Aloy-Prosper A, Penarrocha-Oltra D, Penarrocha-Diago MA, Penarrocha-Diago M. The outcome of intraoral onlay block bone grafts on alveolar ridge augmentations: a systematic review. *Med Oral Patol Oral Cir Bucal* 2015:e251–8. http://dx.doi.org/10.4317/medoral.20194.
34. Aghaloo TL, Moy PK. Which hard tissue augmentation techniques are the most successful in furnishing bony support for implant placement? *Int J Oral Maxillofac Implants* 2007;**22**(Suppl):49–70.
35. Milinkovic I, Cordaro L. Are there specific indications for the different alveolar bone augmentation procedures for implant placement? A systematic review. *Int J Oral Maxillofac Surg* 2014;**43**:606–25. http://dx.doi.org/10.1016/j.ijom.2013.12.004.
36. Keestra JA, Barry O, Jong L, Wahl G. Long-term effects of vertical bone augmentation: a systematic review. *J Appl Oral Sci* 2016;**24**:3–17. http://dx.doi.org/10.1590/1678-775720150357.
37. Camps-Font O, Burgueno-Barris G, Figueiredo R, Jung RE, Gay-Escoda C, Valmaseda-Castellon E. Interventions for dental implant

placement in atrophic edentulous mandibles: vertical bone augmentation and alternative treatments. A meta-analysis of randomized clinical trials. *J Periodontol* 2016;**87**:1444–57. http://dx.doi.org/10.1902/jop.2016.160226.

38. Clementini M, Morlupi A, Canullo L, Agrestini C, Barlattani A. Success rate of dental implants inserted in horizontal and vertical guided bone regenerated areas: a systematic review. *Int J Oral Maxillofac Surg* 2012;**41**:847–52. http://dx.doi.org/10.1016/j.ijom.2012.03.016.
39. Clementini M, Morlupi A, Agrestini C, Ottria L. Success rate of dental implants inserted in autologous bone graft regenerated areas: a systematic review. *Oral Implantol (Rome)* 2011;**4**:3–10.
40. Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med Inform Decis Mak* 2007;**7**:16. http://dx.doi.org/10.1186/1472-6947-7-16.
41. Egger M, Smith GD, Sterne JA. Uses and abuses of meta-analysis. *Clin Med (Lond)* 2001;**1**:478–84.
42. Burda BU, Holmer HK, Norris SL. Limitations of A Measurement Tool to Assess Systematic Reviews (AMSTAR) and suggestions for improvement. *Syst Rev* 2016;**5**:58. http://dx.doi.org/10.1186/s13643-016-0237-1.
43. Perry R, Leach V, Davies P, Penfold C, Ness A, Churchill R. An overview of systematic reviews of complementary and alternative therapies for fibromyalgia using both AMSTAR and ROBIS as quality assessment tools. *Syst Rev* 2017;**6**:97. http://dx.doi.org/10.1186/s13643-017-0487-6.
44. Moraschini V, Barboza Edos S. Quality assessment of systematic reviews on alveolar socket preservation. *Int J Oral Maxillofac Surg* 2016;**45**:1126–34. http://dx.doi.org/10.1016/j.ijom.2016.03.010.
45. De Buitrago JG, Avila-Ortiz G, Elangovan S. Quality assessment of systematic reviews on alveolar ridge preservation. *J Am Dent Assoc* 2013;**144**:1349–57.
46. Chambrone L, Faggion Jr CM, Pannuti CM, Chambrone LA. Evidence-based periodontal plastic surgery: an assessment of quality of systematic reviews in the treatment of recession-type defects. *J Clin Periodontol* 2010;**37**:1110–8. http://dx.doi.org/10.1111/j.1600-051X.2010.01634.x.
47. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009;**62 e**:1–34. http://dx.doi.org/10.1016/j.jclinepi.2009.06.006.
48. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999;**354**:1896–900.
49. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Int J Surg* 2014;**12**:1495–9. http://dx.doi.org/10.1016/j.ijsu.2014.07.013. STROBE Initiative.
50. Lemos CA, Ferro-Alves ML, Okamoto R, Mendonca MR, Pellizzer EP. Short dental implants versus standard dental implants placed in the posterior jaws: a systematic review and meta-analysis. *J Dent* 2016;**47**:8–17. http://dx.doi.org/10.1016/j.jdent.2016.01.005.
51. Garaicoa-Pazmino C, Suarez-Lopez del Amo F, Monje A, Catena A, Ortega-Oller I, Galindo-Moreno P, Wang HL. Influence of crown/implant ratio on marginal bone loss: a systematic review. *J Periodontol* 2014;**85**:1214–21. http://dx.doi.org/10.1902/jop.2014.130615.

Address:
Duarte Nuno da Silva Marques
Faculdade de Medicina Dentária da Universidade de Lisboa
Cidade Universitária
1649-003 Lisboa
Portugal
Tel.: +35 1966486375
E-mail: marques@fmd.ulisboa.pt