Original Article

Characterizing the orthodontic research literature: 2020

David W. Chambers[a]

ABSTRACT

Objectives: To characterize features of the current orthodontic literature.

Methods and Materials: All research articles published in 2020 (N = 350) in the *American Journal of Orthodontics and Dentofacial Orthopedics, The Angle Orthodontist,* and the *European Journal of Orthodontics* were categorized on 48 features such as type of study (domain of generalization, subjects, and research design), analytical tools (statistical tests, power and normality of data, consistency of measurement, management of covariables, and corrections of multiple independent tests), and reporting characteristics. Consistency of the coding was high ($\kappa > .990$).

Results: The "most typical" article was a cohort study reporting multiple patient outcomes at a single treatment location. Soundness of analyses was uneven, with about half providing information on power or normality of the data and consistency of measurement. Few articles addressed covariables or adjusted for multiple tests of independent outcomes. Photos and flow charts were commonly used to explain methods, and results were presented in multiple formats. There was a clear association between design and reporting characteristics and type of study for systematic reviews, meta-analyses, and case reports. There were small but consistent differences across the three journals.

Conclusions: The quality of the orthodontic research literature has advanced at an uneven pace, and this review identifies areas that could be strengthened. Substantial gaps remain in achieving accepted standards for randomized controlled trials and opportunities exist for better understanding measures of effect through design and analysis using regression techniques to identify sources of variance. (*Angle Orthod.* 2023;93:228–235.)

KEY WORDS: Research design; Reporting; Standards; Questions

INTRODUCTION

Orthodontic research has changed significantly in the past 50 years. Ethical standards for studies involving patients became formalized as a result of the Belmont Report in the United States, the Declaration of Helsinki in Europe, and elsewhere in the world. Increased funding from governments and industry prompted more scholarship. Electronic storage and key-word retrieval eased the entry of residents and first-time scholars. A one-way analysis of variance (ANOVA) is no longer considered an advanced statistical test. Computer programs now model latent variables (patterns), and published articles themselves have become dependent variables in systematic reviews (SRs) and meta-analyses (MAs). All of this has been accompanied by publication standards such as Consolidated Standards of Reporting Trials (CONSORT) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) that identify best practices. Documenting key changes is important for strengthening the science base, providing context to reviewers and editors, alerting those who teach residents about needed skills, and signaling to practitioners how researchers are using more sophisticated approaches.

There is literature critiquing various shortcomings in the field. Examples include conflict of interest,[1] retracted articles,[2] conformance with reporting standards such as CONSORT for randomized controlled trials (RCTs)[3–9] and SRs,[10,11] focus group reporting,[12] power (sample size) estimates,[13,14] cohort studies that test against baseline,[15,16] and country of authorship.[17,18]

[a] Professor, Department of Diagnostic Sciences, Arthur A. Dugoni School of Dentistry, University of the Pacific, San Francisco, Calif, USA.

Corresponding author: David W. Chambers, Professor, Department of Diagnostic Sciences, Arthur A. Dugoni School of Dentistry, University of the Pacific, 155 Fifth Street, San Francisco, CA 94103, USA
(e-mail: dchamber@pacific.edu)

Accepted: July 2022. Submitted: April 2022.
Published Online: October 20, 2022

An exhaustive survey was performed of three journals during the year 2020 to quantify the topics, research methods, analytical approaches, and conventions for describing materials and methods and reporting findings. This is an update of previous studies,[3,19] the most recent having covered studies reporting data 7 years previously.

MATERIALS AND METHODS

The sample included all articles published in 2020 in the *American Journal of Orthodontics and Dentofacial Orthopedics* (AJO-DO), *The Angle Orthodontist* (Angle), and the *European Journal of Orthodontics* (EJO). There were 172 articles in AJO-DO (10 of which were in digital-only format), 100 articles in Angle, and 78 articles in EJO.

Coding categories were developed based on review of articles in the June (or second) 2019 issues of each journal. Adjustments were made in the coding template based on reading and coding the August (or fourth) issue in that year. The coding on which this analysis was based was conducted throughout 2020 as each issue became available. All articles were coded twice by the author with the second coding used for analysis.

The coding template is shown in Table 1. Some characteristics were coded in mutually exclusive categories. For example, each study was classified as involving humans, animals, or previously published literature. Where appropriate, multiple categories were marked present, absent, or not appropriate when multiple characteristics could be present. An example would be cases in which more than one statistical test was performed for a study.

Three groups of features were coded for each published article. Studies were grouped by type of research question. This included the nature of the research question (treatment effectiveness, perception or measurement issues, biology, case studies, and SRs, for example), research design (clinical trials, sample descriptions, quantification of variance contributed by various factors, and consistency of measurement approaches, for example), and subjects (patients in treatment, general population, tissues, animals, or passes with phantoms, what was counted in determining sample size). Type of analysis, the second major coding characteristic, included enumeration of the various statistical tests performed. Also counted were the presence of data management practice such as power tests and tests for normality of data, covariable handling, and estimates of consistency of measurement. The third major group of characteristics identified were concerned with reporting. This included use of photos, tables, flow charts, and figures in the presentation of the study design and for presenting findings, proportion of space devoted to various sections of the articles, number of references, and country of authors.

The results of analyses were reported in contingency tables as counts and percentages. Chi-square analyses, or Fisher exact tests in the case of small expected values, were performed. Intrarater consistency was estimated by Cohen's kappa for the two codings of all variables for all articles by the author and by Fleiss's kappa for interrater reliability using a random sample of 20 characteristics for 10 articles coded by four raters.

RESULTS

The data set contained 16,800 data points (48 categories for 350 articles). Interrater consistency was $\kappa = .995$; interrater consistency was $\kappa = .867$.

Combining the most frequently occurring features in the data set, the "typical" orthodontic article published in AJO-DO, Angle, and EJO in 2020 had the following characteristics: a retrospective analysis of patient records at a single site to determine whether any of multiple measured occlusal outcomes differed across treatment. About half of the studies reported power calculations or checks for normality of data, and half did not report the consistency of data capture. The *t*-test for differences between two means was the most common statistical test. A variety of ways of describing the materials and methods and displaying the results was used. Tables of numerical findings were very common. Flow charts and forest plots were strongly associated with SRs or MAs, and before and after photos were ubiquitous in case reports. Together, materials and methods and results comprised 70% of each article, with about 10% devoted to introducing the issues studied. There was an average of nearly 35 references per article. Contributions were published from researchers around the world.

Study Type

As shown in Table 2, almost 60% of articles drew conclusions regarding changes in occlusion attributable to treatment interventions. The smallest category was biology studies, with 5% of the articles. AJO-DO had the largest proportion of articles describing population characteristics. Angle offered proportionally more reports on treatment effects. EJO presented more science regarding the biological characteristics of orthodontics ($\chi^2 = 24.83$, $df = 8$, $P = .002$, $\varphi = 266$).

Table 3 summarizes the results concerning the unit of analysis. More than three-quarters of articles generalized to the individual human, usually a patient. When tissues, animals, mathematical models, or phantoms were used, this was labeled "nonpersons," and these accounted for 145 of the studies (41%). Secondary analyses comprised about 10%. These

Table 1. Coding Categories for Characteristics of Articles Appearing in 2020 in AJO-DO, Angle, and EJO

Type of Study (All Classifications Are Mutually Exclusive in Each Category)	
Domain of generalization: What kinds of practices or decisions do results apply to?	
Treatment effects: Does treatment X make a difference?	
Population characteristics: What proportion of children in a community have condition X?	
Perceptions: Do patients and orthodontists agree on what a beautiful smile looks like?	
Measurement: Are digital landmark locations consistent with human judgment?	
Biology: Is the presence of X cells different following treatment?	
Subjects, unit of analysis: What are the components of N?	
Patients treated at a single site	
Patients treated at multiple sites	
Nontreatment accompaniments of care	
Communities not in treatment	
Tissue	
Animals (primarily rats)	
Mathematical models	
Phantoms	
Studies in the literature	
Research design: How are data organized for inferential tests?	
Cohort: Two or more naturally occurring groups, retrospective	
Matching: Single cohort chose, then subjects as nearly as possible on some points selected, retrospective	
Clinical trial: Prospective comparison of outcomes across created groups, prospective	
Randomized controlled trial: Clinical trial where participants are assigned on randomization of some features, prospective	
Regression: Estimates of variance contributed by various sources	
Systematic reviews/meta-analysis: Analysis of selected published articles/estimation of common effect given selected articles	
Case: Reports of status of individual patient from diagnosis through treatment	
Observation: Description of characteristics in naturally occurring group	
Latent structure: Computerized identification of patterns in data where input or outcomes are not naturally observable entities	
Analysis (all items scored yes/no/NA)	
Power test	Statistical tests
Test for parametric variables	*t*-test for two means
Management of covariables	Nonparametric test for two means
Correction for multiple dependent variables	χ^2
Consistency check on measurement	Correlation coefficient
ICC	One-way ANOVA
Dahlberg	Repeated-measures ANOVA
κ	N-factorial ANOVA
Bland-Altman	Multiple regression
t-test	Latent structure
Percent agreement	Content analysis
Needed but none performed	
Reporting	
Material and methods sections (yes/no/Not Applicable)	Results (yes/no/Not Applicable)
Numerical data regarding sample	Numerical data on outcome variables
Patients, equipment, intervention	Bar graph
Measurement, landmark location	Chart, scatterplot, survival
Flow chart	Box and whiskers
	Before and after photo
	Forest plot
	Bias threat table
	Chart data for cases
Inches of text	Country of authors
Introduction	United States, Canada
Materials and methods	Europe, Australia
Results, including tables and figures	Asia
Discussion, limits, conclusion	Middle East
	South America
Number of references	

were SRs and MAs where previously published articles were the unit of analysis. EJO reported fewer studies with patients as the unit of analysis and more secondary work ($\chi^2 = 17.47$, $df = 4$, $P = .001$, $\varphi = .223$).

Table 4 displays studies by research design. Retrospective studies, at 32%, and studies using latent structure or regression designed to explain observed outcomes, were used in 28% of publications. Prospec-

Table 2. Domain of Generalization in Articles Published in 2020 in Three Orthodontic Journals[a]

	AJO-DO	Angle	EJO	Total
Treatment effects	56.4	65.0	57.7	59.1
Population characteristics	15.7	7.0	7.7	11.4
Perceptions	14.5	8.0	12.8	12.3
Measurement features	12.2	15.0	9.0	12.3
Biology	1.2	5.0	12.8	4.9
N	172	100	78	350

[a] Values are percentages of articles per journal in each category.

tive and descriptive studies of naturally occurring populations each accounted for about 20%. AJO-DO was less likely to publish research with prospective designs and Angle less likely to publish observational studies. EJO was less apt to publish observational or retrospective research. It did not publish any case studies in 2020 ($\chi^2 = 36.49$, $df = 6$, $P < .001$, $\varphi = .323$).

As shown in Figure 1, RCTs to test interventions to improve patient occlusion accounted for 9% of the published articles. Prospective designs were well represented (about one-third of studies) but used about twice as often for research with animals, modeling, tissue, or phantoms ($\chi^2 = 20.96$, $df = 1$, $P < .001$, $\varphi = .060$).

Study Analysis

Integrity of the data sets and statistical tests are shown in Table 5. Power tests and tests for parametric properties of data appeared in fewer than half of the studies where these would have been appropriate. The effect of covariables (such as age and sex of patients) on tested outcomes was considered in few cases and tested in only 7%. One-third of the studies had multiple dependent variables. This was especially true for cohort studies at 55%. Almost no studies reported make appropriate corrections. The Bonferroni adjustment is simple, but conservative. Multiple analysis of variance (MANOVA) is preferred when it can be performed.

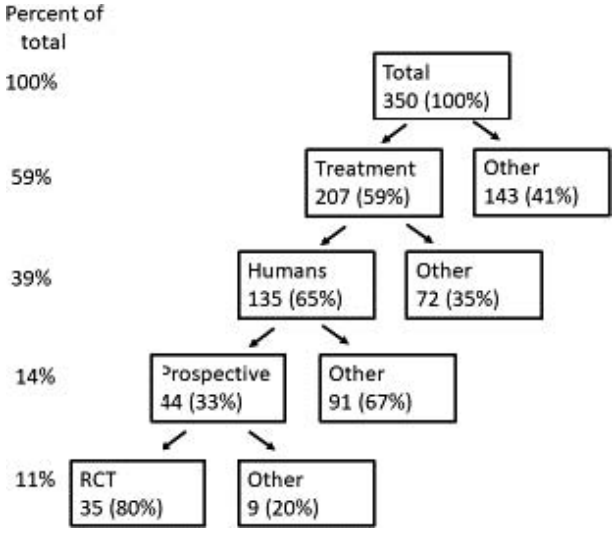

Figure 1. Breakdown of randomized controlled trials by type of conclusions and subject.

Parametric and nonparametric tests for differences between two means were used in almost half of the reported work. Other common tests were chi-square, correlation coefficient, one-way ANOVA, and regression analysis. Less common were tests for complex experimental designs such as repeated-measures ANOVA, N-factorial ANOVA, and latent structure methods.

Study Reporting

Use of graphic and tabular presentations of information in the materials and methods section and results section of articles are displayed in Table 6. In almost half of the studies, diagrams were used to help readers understand landmark and measurement locations. In about a quarter each, flow charts and pictures of equipment were presented. Consistent with the underreporting of covariables, data on characteristics of the sample other than independent and dependent variables were presented in fewer than 10% of cases. Tables showing average and standard deviation

Table 3. Subjects, Units of Analysis[a]

	AJO-DO	Angle	EJO	Total		AJO-DO	Angle	EJO	Total
Patients, single site	53.5	60.0	32.1	50.6	Individuals	81.9	76.0	64.1	76.2
Patients, multiple sites	5.8	3.0	24.4	9.1					
Patients, nontreatment	8.1	8.0	3.8	7.1					
Communities	14.5	5.0	3.8	9.4					
Tissue	1.2	4.0	2.6	2.3	Nonperson	11.6	18.0	17.8	14.3
Rats, dogs	3.5	2.0	9.0	4.3					
Mathematical models	1.7	10.0	2.6	4.3					
Phantoms	5.2	2.0	1.3	3.4					
Literature	6.4	6.0	20.5	9.4	Secondary	6.4	6.0	20.5	9.4
N	172	100	78	350					

[a] Values are percentages of articles per journal in each category.

Table 4. Research Design[a]

	AJO-DO	Angle	EJO	Total		AJO-DO	Angle	EJO	Total
Observation	25.5	13.0	9.0	18.3	Observation	25.5	13.0	9.0	18.3
Cohort	13.4	28.0	16.7	18.3	Retrospective	33.8	36.0	21.8	31.7
Matching	4.7	2.0	5.1	4.0					
Case	15.7	6.0	0.0	9.4					
RCT	11.0	25.0	30.8	19.4	Prospective	13.3	30.0	32.1	22.3
Clinical trial	2.3	5.0	1.3	2.9					
Regression	20.3	15.0	17.9	18.3	Explanatory	27.2	21.0	37.1	27.8
Reviews, meta-analysis	5.2	6.0	19.2	8.6					
Latent structure	1.7	0.0	0.0	0.9					
N	172	100	78	350					

[a] Values are percentages of articles per journal in each category.

scores for various classifications of the outcomes were common in every applicable case (80% overall). These were supplemented with bar graphs, scatterplots, and sometimes box-and-whisker diagrams where tests were performed on differences between group averages or associations among variables. The second most common way of reporting outcomes (about one-third of the cases) was by using photographs. The expected use of data reporting associated with type of project was observed. For example, forest plots always appeared for MAs, bias tables were very common for SRs, and before and after photos and patient chart information always accompanied case reports.

Several demographics features of the literature were also recorded. A total of 1025 author names appeared on the 350 publications. These were categorized in five geographic regions. Almost one-third (31%) of authors were from Europe, followed by Asia (21%), the United States and Canada (18%), Latin America (16%, almost entirely from Brazil), and the Middle East (14%, mostly Turkey and a few from India). EJO featured more articles from European authors ($\chi^2 = 24.32$, $df = 4$, $P < .001$, $\varphi = .069$).

All three publications used a two-column, 8.5×11 format so it was possible to compare allocation of text to sections, measured in inches. The averages were 10.1 column inches for the introduction; 45.5 for materials and methods; 48.7 for results; and 28.3 for discussion, limitations, and conclusions. Graphics were measured in inches in their appropriate sections. The typical article had 33.4 references. The shortest discussion sections (averaging 29.0 inches) were found in EJO ($\chi^2 = 31.24$, $P < .001$), and the fewest references (averaging 28.0) were found in Angle ($\chi^2 = 19.30$, $df = 2$, $P < .001$, $\varphi = 055$).

DISCUSSION

This survey demonstrated the richness of the orthodontic literature. The range of issues addressed and the research methods used to understand them are reflective of a profession that is evolving and has not yet settled into rituals of inquiry. "Definitive experiments," those intended to establish irrefutable, "yes/no" answers to research questions were rare. For the most part, we are looking over the shoulders of practitioners in academic settings who are showing us what they have been exploring. The refinements of power calculations, management of covariables, and adjustments for multiple hypothesis testing are unevenly applied. Several decades ago, these were not matters for consideration. The sophistication of statistical tests also shows strong recent advances. Orthodontic research is now international.

Table 5. Design Features and Statistics[a]

Integrity of the Data		Statistical Tests	
Power test performed	42.6	*t*-test for means	27.7
Test for parametric variables	41.8	Nonparametric for means	18.9
Management of covariables	16.5	χ^2	15.7
Multiple dependent variables	1.1	Correlation coefficient	15.1
Measurement consistency	51.7	One-way ANOVA	18.0
ICC	33.3	Repeated-measures ANOVA	04.6
Dahlberg	13.9	N-factorial ANOVA	07.1
κ	9.0	Multiple regression	16.0
Bland-Altman	.6		
t-test	5.2	Latent structure	02.7
Percentage agreement	.7	Content analysis	00.6

[a] Values are percentages of those cases where such tests are appropriate.

Table 6. Reporting

Materials and Methods		Results		Unusual Value[a]
Characterization of sample	7.7	Numerical tables	78.9	69.8 AJO-DO***
Equipment, intervention	25.1	Bar graphs	20.9	33.8 RCTs***
Measurement, landmarks	47.7	Charts, scatterplots	24.9	34.4 observations***
Flow charts	27.7	Box and whiskers	5.4	
		Photos, before/after	31.7	20.5 EJO**; 78.8 cases***
		Forest plots	4.0	9.0 EJO*; 43.3 SRs***
		Bias threats table	4.0	46.7 SRs***
		Chart data for cases	7.1	75.8 cases***

[a] "Unusual Value" identifies cases where a statistically significant result was found using the χ^2 analysis. The unusual value shows cases where the observed and expected values differ widely.

* $P < .05$; ** $P < 01$; *** $P < .001$.

Strengthening of Standards

Study standards are beginning to appear regularly, such as CONSORT for reporting RCTs; STROBE for observational studies; COREQ for qualitative studies; PRISMA for reporting SRs; and Grading of Recommendations Assessment, Development, and Evaluation and AMSTAR for critiquing them. But these were not as widely evident as were forest plots for MAs and before-and-after images for case presentations.

Literature About the Literature

The findings in this summary were generally consistent with the articles exploring individual features in the literature. Baumgartner et al.[19] also studied AJO-DO, Angle, and EJO, retrospectively, prior to 2012 and found that AJO-DO reported cases but Angle and EJO carried more "research" articles. Bearn and Alharbi[3] reported a different order for publishing research articles in the period 2008 through 2012, with Angle being the least scholarly. Pandis and colleagues[9] presented an overview article in the *Journal of Dental Research* characterizing "publication waste" as "insufficient information to make clear decisions." They found adequate sample size calculations in no more than 36% of studied articles, adequate randomization in no more than 68%, and management of covariables in fewer than 43% of studies. Koletsi et al.[13] reported that sample sizes were inadequate in 71% of the studies they reviewed. Seehra et al.[14] found adequate sample size calculations in only 52% of the articles. In the current study, power tests were performed in 43% of the cases, and management of covariables was observed in 17%. Seehra et al.[14] also reported that 52% of studies of treatment outcomes used only single-site data, 23% were prospective, 19% retrospective, and 53% used cohort designs. In the current article, these numbers were 83%, 63%, 27%, and 37%. Donatelli et al.[20] reported that 47% of the articles they reviewed in five orthodontic journals reported consistency of measurement, with intraclass correlation coefficient (ICC) being the most common method. In this research, 52% of articles reported measurement consistency. The distribution of methods was comparable. Satisfaction with the quality of RCTs and summary reporting in orthodontics generally tended to group in the mediocre range, with slightly less than half of articles meeting reviewers' criteria.[5–9] A recent trend called "RCT spin" has been identified where authors claim in the materials and methods section to have used high levels of research design that cannot be confirmed in the results section. Flint and Harrison[5] placed the proportion of "spin RCTs" at 42%. Harrison[7] put it at 55%. The review article by Koletsi et al.[21] in the *Journal of Evidence-Based Practice* reported that only 40% of claimed RCTs they reviewed met the criteria. In the current study, articles were classified as RCT if the authors labeled them as such, but studies claiming to randomly assign subjects to treatment conditions were very rare. Chen et al.[4] reviewed five leading orthodontic journals in 2019 and reported that the top level of evidence, using the GRADE system, was achieved in only 1% of articles analyzed.

Opening the Question

Recent attention in the secondary literature focused on strengthening the RCT model to enhance protection of claimed treatment benefits from bias. Claims for differences between groups were the most common study reported in the literature sampled. However, these were usually retrospective and performed at a single location, perhaps with a single operator or common protocol for patient selection. Analysis of variance tests (including the *t*-test and its nonparametric versions) were the most popular form of statistical analysis, and often, multiple such tests were applied to a single study. Although calls for improving this model are appropriate, the CONSORT standard is actually silent on this approach. Instead, it is urged that measures of effect be reported to quantify the anticipated impact of intervention on outcomes of interest.

More than a quarter (28%) of the 2020 articles studied used a regression approach intended to

measure effect and to identify which of several candidate sources of variance mattered most to producing the observed outcomes. More regular use of regression approaches would have the double advantage of addressing concerns over covariables and increasing the prominence of measures of effect, such as proportion of explained variance. A full enumeration of the methods of analysis suggested a trend toward more elaborate analysis, including latent stature approaches that were not even possible a few decades ago.

The orthodontic research literature is diverse and representative of the range of inquiry now being used to build stronger practices. Although conducted primarily by residents under the tutelage of university faculty[22] and responsive to those dedicated to evidence-based dentistry, it should retain its dynamic nature. It has become truly international[17,18] and represents a diversity of topics and approaches. It is easy enough to find examples of design and methodology that could be improved or at least standardized, but that, in itself, is not sufficient. What is needed is scholarship that deepens our understanding of theory and promotes better practice. This involves both high standards and a large pipeline. The true guardians and promoters of quality are the reviewers and editors who identify articles that address the most important topics and expect rigor that will reduce the potential for plausible unexplained causes.

CONCLUSIONS

- The orthodontic literature is diverse with respect to topic and intended contributions, research design and analysis, and reporting standards.
- Emerging high standards associated with RCTs are not widely applied.
- Important topics may not lend themselves to RCT approaches, and new methods grounded in measures of effect and identification of sources of variance are appearing.
- It is important to periodically benchmark the literature to provide insight into changes in the field as well as requirements for upgrading training of residents and practitioners who read the literature.

REFERENCES

1. Eliades T, Turpin DL. Conflict of interest: always report it, and if in doubt, ask. *Amer J Orthod Dentofac Orthoped*. 2008;134(3):327–328.
2. Rapani A, Lombardi T, Berten F, Del Lupo V, Di Lenardo R, Stacchi C. Retracted publication and their citation in dental literature: a systematic review. *Clin Exper Dent Res*. 2020; 6(4):383–390.
3. Bearn DR, Alharbi F. Reporting of clinical trials in the orthodontic literation from 2008 to 2012: observational study of published reports in four major journals. *J Orthod*. 2015; 42(3):180–191.
4. Chen Y, Hua F, Mei Y, Thiruvenkatachun B, Riley P, He H. Characteristics and level of evidence of clinical studies published in 5 leading orthodontic journals. *J Evid Based Pract*. 2019;19(3):273–282.
5. Flint HE, Harrison JE. How well do reports of clinical trials in the orthodontic literature comply with the CONSORT statement? *J Orthod*. 2010;37(4):250–261.
6. Gibson R, Harrison J. What are we reading? An analysis of orthodontic literature 1998-2008. *Amer J Orthod Dentofac Orthoped*. 2011;139(5):e471–e484.
7. Harrison JE. Clinical trials in orthodontics II: assessment of the quality of reporting of clinical trials published in three orthodontic journals between 1989-1998. *J Orthod*. 2003; 30(4):309–315.
8. Hua F, Deng L, Kau CH, Jiang H, He H, Walsh T. Reporting quality of randomized controlled trail abstracts: survey of leading general dental journals. *J Amer Dent Assoc*. 2015; 146(9):669–678.
9. Pandis N, Fleming PS, Katsaros C, Ioanaidis JPA. Dental research waste in design, analysis, and reporting: a scoping review. *J Dent Res*. 2021:100(3):245–252.
10. Papageorgiou SN, Papadopoulos MA, Athanasiou AE. Evaluation of methodology and quality characteristics of systematic reviews in orthodontics. *Orthod Craniofac Res*. 2011;14(3):116–137.
11. Vasquez-Cardonas J, Zapata-Noreña O, Carevjal-Florez A, Barbosa-Liz DM, Giannakoulos NN, Farrian CM. Systematic reviews in orthodontics: impact of the PRISMA for Abstracts checklist on completeness of reporting. *Amer J Orthod Dentofac Orthoped*. 2019;156(4):442–452.
12. Al-Moghrabi D, Tsichlaki A, Al Kadi S, Fleming PS. How well are dental qualitative studies involving interviews and focus groups reported? *J Dent*. 2019;84:44–48.
13. Koletsi D, Fleming PS, Seehra J, Bagos PG, Pandis M. Are sample sizes clear and justified in RTCs publications in dental journals? *PLOS ONE*. 2014, DOI 10.13711/journal.pone.0085949.
14. Seehra J, Stonehouse-Smith D, Cobourne MT, Tsagris M, Pandis N. Are treatment effects assumptions in orthodontic trials overoptimistic? *Eur J Orthod*. 2021;43(5):583–587.
15. Gratsia S, Koletsi D, Fleming PS, Pandis N. Statistical testing against baseline in orthodontic research: a meta-epidemiologic study. *Eur J Orthod*. 2019;41(2):165–171.
16. Wadgave U, Khairaur MR, Kadu TS, Wadlgave Y. Evaluation of dental trails comparing baseline differences using p values. *Acta Odontol Scand*. 2019;77(3):181–183.
17. Alquaydi AR, Kanavakis G, Nasser-Ud-Din S, Athanasiou A. Authorship characteristics of orthodontic randomized controlled trials, systematic reviews, and meta-analysis in non-orthodontic journals with impact factors. *Eur J Orthod*. 2018; 40(5):480–487.
18. Kanavakis G, Spinos P, Polychronpoulou A, Eliades T, Papadopoulou MA, Alhanasiou AE. Orthodontic journals with impact factors in perspective trends in the types of articles and authorship characteristics. *Amer J Orthod Dentofac Orthoped*. 2006;130(4):516–522.
19. Baumgartner S, Pandis N, Eliades T. Exploring the publications in three orthodontic journals: a comparative

analysis of two 5-year period. *Angle Orthod*. 2012;84(3): 397–403.

20. Donatelli RE, Park J-A, Abdullah Alghamdi YM, Pandis N, Lee S-J. Assessment of reliability in orthodontic literature: a meta-epidemiological study. *Angle Orthod*. 2022;92(3):409–414.
21. Koletsi D, Pandis N, Polychronopoulou A, Eliades T. Mislabeling controlled clinical trials (CCTs) as "randomized clinical trials (RCTs)" in dental specialty journals. *J Evid Based Pract*. 2012;12(3):124–130.
22. Chambers DW. Does our research tool kit equip us to make generalizable claims about dental education? *Eur J Dent Educ*. 2012;16:202–207.