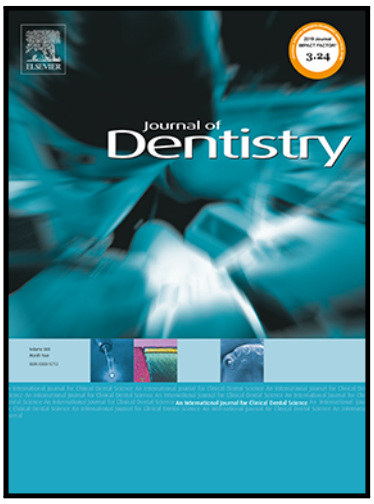
Journal Pre-proof



Are longitudinal randomised controlled oral health trials properly analysed? A meta-epidemiological study

Samer Mheissen , Haris Khan , Jadbinder Seehra , Nikolaos Pandis

PII: S0300-5712(22)00238-X
DOI: https://doi.org/10.1016/j.jdent.2022.104182
Reference: JJOD 104182

To appear in: *Journal of Dentistry*

Received date: 27 March 2022
Revised date: 6 June 2022
Accepted date: 8 June 2022

Please cite this article as: Samer Mheissen , Haris Khan , Jadbinder Seehra , Nikolaos Pandis , Are longitudinal randomised controlled oral health trials properly analysed? A meta-epidemiological study, *Journal of Dentistry* (2022), doi: https://doi.org/10.1016/j.jdent.2022.104182

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 Published by Elsevier Ltd.

source: https://doi.org/10.48350/170634 | downloaded: 12.10.2024

Journal Pre-proof

Are longitudinal randomised controlled oral health trials properly analysed? A meta-epidemiological study

Title page

Are longitudinal randomised controlled oral health trials properly analysed? A meta-epidemiological study

- Samer Mheissen, DDS, Syrian Board in Orthodontics, Former instructor in Orthodontic Department, Syrian Ministry of Health, Specialist Orthodontist in Private Practice, Syria- Damascus
- Haris Khan Associate Professor of orthodontics, CMH institute of dentistry Lahore, National University of Medical Sciences, Punjab, Pakistan.
- Jadbinder Seehra, Senior Clinical Teacher and Consultant Orthodontist, Centre for Craniofacial Development & Regeneration, Faculty of Dentistry, Oral & Craniofacial Sciences, King's College London, Floor 25, Guy's Hospital, London, SE1 9RT, United Kingdom.
- Nikolaos Pandis, Department of Orthodontics and Dentofacial Orthopedics, Dental School/Medical Faculty, University of Bern, Switzerland
-

Running title: Longitudinal Dental RCTs analyses

Keywords, Longitudinal data, statistical analysis, clinical trials, dental, oral health

Corresponding Author:

Samer Mheissen, Former instructor in Orthodontic Department, Syrian Ministry of Health, Private Practice, Damascus, Syria.

Email: Mheissen@yahoo.com

Word count: 2775

Reference count: 32

Figure/table counts: 2 figures and 2 tables

Author contributions

Samer Mheissen: Contributed to conception, design, data acquisition and statistical analyses, drafted and critically revised the manuscript. Haris Khan: Contributed to conception, design, data acquisition, and critically revised the manuscript. Jadbinder Seehra: Contributed to conception, design, interpretation, drafted and critically revised the manuscript. Nikolaos Pandis:

Journal Pre-proof

Contributed to conception, design, data acquisition and statistical analyses, drafted and critically revised the manuscript. All authors gave their final approval and agree to be accountable for all aspects of the work.

Abstract

Introduction

Longitudinal designs with multiple outcome measurements are commonly encountered in oral health randomised controlled trials (RCTs). The aim of this meta epidemiological study was to assess whether optimal statistical analysis approaches have been used in longitudinal oral health RCTs.

Data sources

PubMed search was undertaken in September 2021 for longitudinal oral health RCTs with at least 3 repeated outcome measurements which have been published between 2016- 2020 in the highest impact general and specialty dental journals.

Study selection

Study selection and data extraction were accomplished independently and in duplicate. The statistical methods undertaken in the selected articles were tabulated, and the association between study characteristics and use of optimal analyses were assessed using X^2 or Fisher's exact test and logistic regression.

Results

Five hundred and five oral health RCTs were deemed eligible for inclusion. Of these, only 28.3% RCTs used optimal statistical analyses for a longitudinal trial design. For the trials with an optimal statistical approach, the most frequent test used was repeated measures analysis of variance (RM-ANOVA) followed by mixed effect models (MEM). The use of optimal statistical

Journal Pre-proof

tests was predicated by the involvement of a statistician (OR: 2, 95% CI:1.27 - 3.18, p<0.01), the journal impact factor (OR:1.19, 95% CI;1.1 - 1.29), continent of first author (likelihood ratio test p=0.01), number of the authors (OR:1.22 , 95% CI;1.12-1.3, p<0.001), protocol registration (OR: 1.48, 95%CI; 1 to 2.2, p=0.05), funding(OR:2.4, 95%CI; 1.6 - 3.7, p<0.001), and dental specialty (likelihood ratio test p<0.001).

Conclusions

Most longitudinal oral health RCTs did not use optimal statistical analyses. Greater awareness of optimal analyses used to assess longitudinal data reported in oral health trials is required to circumvent the reporting of suboptimal inferences, selective reporting and research waste.

Clinical significance

Further progress is required to avoid suboptimal statistical analyses and fully utilise the benefits of the repeated measurements over time in oral health RCTs.

Keywords

Longitudinal data, statistical analysis, clinical trials, dental, oral health

Journal Pre-proof

Introduction

As the cornerstone of healthcare evidence, clinical trials have been instrumental to major healthcare advances. Although randomised controlled trials (RCTs) represent the gold standard to detect the efficacy and harms of an intervention, meta-epidemiological studies in dentistry have identified issues with RCTs related to ignoring clustering [1, 2], confounding [3], and data dependencies [4] in the analysis of clinical trials. In addition, incorrect comparisons against baseline values [5, 6], and multiple comparisons are common practices that can lead to incorrect interpretation and biased results [7, 8].

Several trials in dentistry include repeated measurements over time. The assessment of the intervention effect and its possible change over time is an important and relevant clinical question in dentistry. An example of this would be measuring post-surgical pain by collecting outcome data at multiple or more subsequent days instead of at a single time point. Statistical analyses involving comparisons between treatment arms at all time points, comparisons of mean effects over time between groups and comparisons within groups could be undertaken which, however, are considered suboptimal approaches because they ignore the data structure and do not fully utilise the benefits of the repeated measurements over time. For instance, if comparisons at all time points were undertaken then a separate analysis would be required at each time point [9-11]. Such an approach can potentially result in incorrect P-values, incorrect interpretation of the results, loss of power and research waste [12, 13]. In addition, multiple tests increase the probability of false-positive results [14-16] and encourage selective reporting of only significant results [17]. Loss of information and loss of study power is a consequence of using mean effects because again the structure of the data is ignored. The within groups statistical testing also does not give an answer to the question at hand as it is incorrect and

Journal Pre-proof

misleading [18]. This approach ignores the data dependencies and time/placebo effects that can be expressed as improvements in the control group due to time and/or due to the fact that patients are receiving a treatment in a trial. Such an improvement in the control group is not considered and the difference in the outcome between arms is falsely magnified [5]. Unfortunately, this incorrect practice is common in the orthodontic [6] and dental literature [5]. In contrast, longitudinal data analysis could be performed which would be considered the most optimal approach in this situation [9-11].

A more optimal approach for analysis of longitudinal designs is to apply longitudinal statistical methods. The benefit of this approach is that it considers data structures and correlations arising from the multiple observations nested within the same patient, which helps to distinguish between changes within and between patients [10, 11]. Furthermore, longitudinal data analysis increases study power [19] and accounts more efficiently for missing data [20]. A limitation of this analysis is that it requires advanced training in statistical theory and methodology [10, 11, 21, 22], which may explain why it does not appear to be routinely used in oral health [4].

Within the literature there appears to be some indication that the analysis of correlated and longitudinal data in dentistry, appears to be suboptimal[1, 4]. However, a wide-ranging assessment of the issue in oral health trials is lacking. The aim of this investigation was to highlight the methods used for the analysis of longitudinal data and to determine on an exploratory basis any association between study characteristics and the appropriateness of the chosen approaches in oral health randomised trials.

Journal Pre-proof

Material and methods

Eligibility criteria

A selection of dental speciality oral health journals with the highest impact factor (2020) were included in this study. RCTs published between 1st January 2016 and 31st December 2020 in the following journals were sourced: Journal of Dental research, Journal of American Dental Association, European Journal of Orthodontics, American Journal of Orthodontics and Dentofacial Orthopaedics, Angle Orthodontist, Journal of Clinical Periodontology, Journal of Periodontology, Journal of Endodontics, International Journal of Oral and maxillofacial surgery, Journal of Oral and Maxillofacial Surgery, Paediatric dentistry, European Journal of Paediatric dentistry, International Endodontic Journal, Journal of Prosthetic Dentistry, Journal of Prosthodontic Research, and Journal of Prosthodontics.

The terms ‘‘randomised controlled trial’’ was screened in the title, abstract and methodology of the article. As per the Cochrane criteria for the identification of RCTs the following inclusion criteria was applied: human participants, interventions related to healthcare, experimental studies, presence of a control or comparative group, randomisation of participants to control and treatment groups, other trials with terminology in the title or abstract such as ‘‘prospective’’, ‘‘comparative’’, ‘‘efficacy’’ or where an indication was given that a comparison of treatment groups was undertaken prospectively. Additionally, randomised controlled trials (RCTs) of longitudinal design (with 3 or more outcome collection time-points) and with 2 or more treatment arms from the aforementioned journals were included. Observational, animal, preclinical studies and studies in a language other than English were excluded.

Search and selection of studies

An electronic database search was undertaken using Medline via PubMed (www.pubmed.ncbi.nlm.nih.gov) by one author (SM) in September 2021. A search filter of a

Journal Pre-proof

randomised controlled trial published between 1st January of 2016 and 31st December 2020 was employed (Supplemental Table1). All titles and abstracts were screened independently by two authors (SM and HK). The full article text was retrieved when eligibility was either met or unclear.

Data extraction

Two authors (SM and HK) analysed the full text of the remaining records against the inclusion criteria and extracted data independently. Any disagreements were resolved by discussion among the authors. The appropriateness of the statistical analysis was rechecked by another author (NP) with expertise in statistical methods. A pilot assessment of ten random RCTs was undertaken between two authors (SM and HK) to ensure consistency in data extraction variables. A high level of agreement was established. All data was entered into a pre-piloted Microsoft Excel® (Microsoft, Redmond, WA) data collection sheet. Study characteristics from each RCT were extracted as follows: journal name, journal specialty (general dentistry, endodontics, orthodontics, pedodontics, periodontics, prosthodontics, oral surgery), journal impact factor, year of publication, the continent of the first author (Europe, Americas, Asia and other), number of authors, study design (multi-centre or single-centre, split-mouth, parallel, or cross over), number of treatment arms, number of outcome collection time points, the statistical analyses used for the between-group comparisons, involvement of statistician (no or yes; inferred from materials and methods section or authors' acknowledgement), study protocol registration (no/unclear or yes), funding(no/unclear or yes), and significance of results based on the primary outcome. In the absence of no clear primary outcome, the first outcome was analysed: significant or non-significant.

Statistical analysis

Journal Pre-proof

Descriptive statistics were performed. Associations between study characteristics and the statistical methods appropriateness were assessed using X^2 or Fisher's exact test. A logistic regression model was applied for the significant predictors. All statistical analyses were conducted using Stata 15.1 (Stata Corp, TX, USA) and R Software version 4.1.2 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Following application of the eligibility criteria, five hundred and five oral health RCTs were included for full data extraction (Figure 1). Within this sample, 80.8% were of parallel design while 16.6% were split-mouth designs. The median number of time points was 4 (range of 3-40). The majority of the trials (69.1%) reported a 1:1 treatment arm allocation ratio. The median number of authors was 5 (range 1-20), and the median of impact factor of the included journals was 4.17 (IQR 4.34) with a range of (1.87 - 8.73) (Table 1). Only 28.3% of trials used an optimal statistical analysis with an association between the appropriateness of the analysis and the continent of the first author, the journal specialty, the involvement of a statistician, registration, funding, and the statistical significance of the results evident (Table 1). The most used optimal analyses in oral health longitudinal trials were repeated measured analysis of variance (RM-ANOVA) and mixed effect models (MEM), while the t test and its equivalent non-parametric test were the most prevalent suboptimal test (Table 2).

The odds of a significant result were higher for studies with suboptimal analyses compared to studies with optimal analyses (OR: 3.3, 95% CI: 2.14 - 5.1, p<0.01). The involvement of a statistician was associated with higher odds of the use of an optimal analysis compared to suboptimal analyses (OR: 2, 95% CI:1.27 - 3.18, p<0.01). Journals with higher impact factors have higher odds of the use of an optimal analysis compared to suboptimal analyses (OR:1.19, 95% CI;1.1 - 1.29). Optimal/suboptimal analysis was associated with continent (likelihood ratio test p=0.01), authors' number (OR:1.22, 95% CI;1.12-1.3, p<0.001)

Journal Pre-proof

and specialty (likelihood ratio test p<0.001). Likewise, the registration of the prior protocol was associated with a higher odd of optimal analysis compared to non-registered trials (OR: 1.48, 95%CI; 1 to 2.2, p=0.047). Funded studies were associated with more optimal analysis(OR:2.4, 95%CI; 1.6 - 3.7, p<0.001). (Figure 2, Supplemental Table 2)

Discussion

It is not uncommon to encounter longitudinal trial study designs being employed in the assessment of the efficacy of dental interventions which is supported by the fact that from the one thousand and sixty-four screened records five hundred and five oral health RCTs were identified. However, it is important to understand that a longitudinal study design is characterised by repeated measurements/observations resulting in the generation of outcomes from the same patient at multiple time points. On this basis, the observations are then considered to be correlated [23]. Failure to recognise and to account for this characteristic in the statistical analyses could result in both incorrect inferences and biased conclusions[21, 24, 25]. Despite the reported benefits of longitudinal data analysis [10], it appears to be infrequently undertaken within the dental specialities[4]. The results of the current investigation mirrors those of previous studies [1, 4] with just over 70% of oral health longitudinal trials undertaking suboptimal analyses which assume that the measurements are independent rather than assuming they are correlated.

Associations were identified between significance of results, continent, specialty, funding, trial registration, involvement of a statistician and journal impact factor and the

Journal Pre-proof

appropriateness of the analysis used (optimal vs sub-optimal). It would appear that this is counter- intuitive, that using a sub-optimal analysis to analyse the data results in a significant result. However, this could be a reflection of the use of multiple tests that can result in the multiplicity problem where the probability of detecting a false-positive result increases above the conventional 5% [14-16] . The reporting of significant results when sub-optimal analyses are applied may imply publication bias, loss of information [19], possibly mishandling of missing data [20], selective reporting [17] and HARking (hypotheses are formed after looking at the results) [26, 27].

Including a statistician increased the odds of optimal analysis being undertaken. Within the literature, the inclusion of a statistician has been reported to result in better reporting quality of dental trials [28-30]. Although, the latter may not clarify if the appropriate analysis has been undertaken, it does highlight that seeking the appropriate support and guidance from a statistician during the trial planning/analysis may limit the conduct and reporting of inappropriate analyses. This is imperative given the fact that the handling of longitudinal data requires further training in statistical theory and methodology [10, 11, 21, 31]. Incidentally, there was a lack of a clear description of the statistical methods used in most included oral health trials. This is further compounded by the fact that the involvement of a statistical advisor/statistician in the design/analysis was reported only in 1 out of 5 included oral health trials RCTs. Despite the lack of clear instructions pertaining to longitudinal data, to improve the transparency of the statistical methods used, it is recommended that authors follow the SAMPL guideline on statistical methodology [32].

The journal specialty has an impact on using optimal or suboptimal analysis. Endodontic and maxillofacial journals have used more often less optimal approaches when compared with the other journals. The periodontic, orthodontic, and general dental journals were more likely to use optimal approaches in analysing longitudinal data. Higher impact factor was associated with the use of optimal analysis methods, even in the same journal specialty (Supplemental Table 2)

Journal Pre-proof

Similar to previous findings [4], the most frequent of the optimal analyses used to analyse longitudinal data, were repeated-measures analysis of variance (RM-ANOVA), followed by mixed-effect models (MEM). Analyses less frequently used to analyse correlated data in oral health trials included the Friedman test or Generalized Estimating Equations (GEE). Several oral health trials (50.27%) used the t-test and ANOVA or their non-parametric equivalents per time point for inferential statistics, thus failing to recognise and take advantage of the data structure. Possible reasons for the lack of use of appropriate statistical tests may include: a lack of awareness and training on the most appropriate statistical tests, a fixation on finding statistical significance and HARking where hypotheses are formed after looking at the results [26, 27]. It seems that the peer-review process does not hinder the publication of studies with less optimal statistical testing. Typically, a separate statistical review is not required/not available, and the reviewer/editor have to rely on his/her own statistical knowledge. Additionally, user friendly and accessible statistical software packages have made it less dauting for clinicians to analyse their own data. The downside to this is that clinicians may fail to understand the theory behind the analyses and its suitability for the data.

Finally, and in the context of transparency and better use of the available evidence it is also recommended for authors of all studies to share anonymised raw data and this should not exclude longitudinal studies.

Only a single electronic database was searched in this investigation which may mean potentially relevant trials were not identified but including only definite PubMed indexed journals may overcome this issue in the present study. Five hundred and five studies were assessed, and this should be considered a large enough sample to ascertain a baseline of the issue of suboptimal use of analyses in the analysis of longitudinal data. The overwhelming evidence against the use optimal analysis methods for longitudinal data is unlikely to change the key message of this study with the addition of any trials that may have been missed. Within the

Journal Pre-proof

study methodology steps were taken to further minimize bias. These include independent and duplicate screening of studies and pre-piloting prior to data extraction. These were also undertaken under the supervision of an experienced investigator/statistician with experience in analysing longitudinal data.

Conclusions

Prospective oral health randomised longitudinal trial designs are commonly undertaken. In this sample, only 28.3% longitudinal clinical trials used optimal statistical analysis approaches to analyse the data. The most frequent appropriate analyses used were RM-ANOVA and MEM analyses. The involvement of a statistician, specialty, continent, funding, registration and journal impact factor were associated with optimal analysis of longitudinal data. Greater awareness of optimal analyses used to assess longitudinal data reported in oral health trials is required to circumvent the reporting of inappropriate inferences, conclusions, and research waste.

Competing interests

The authors of this study have no conflict of interest in the study.

Funding

No funding was received for this project.

Acknowledgements

None

Journal Pre-proof

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] P.S. Fleming, D. Koletsi, A. Polychronopoulou, T. Eliades, N. Pandis, Are clustering effects accounted for in statistical analysis in leading dental specialty journals?, Journal of dentistry 41(3) (2013) 265-70.

[2] D. Koletsi, N. Pandis, A. Polychronopoulou, T. Eliades, Does published orthodontic research account for clustering effects during statistical data analysis?, European Journal of Orthodontics 34(3) (2012) 287-92.

[3] A. Spanou, D. Koletsi, P.S. Fleming, A. Polychronopoulou, N. Pandis, Statistical analysis in orthodontic journals: are we ignoring confounding?, European Journal of Orthodics 38(1) (2016) 32-38.

[4] S. Mheissen, H. Khan, M. Almuzian, E.E. Alzoubi, N. Pandis, Do longitudinal orthodontic trials use appropriate statistical analyses? A meta-epidemiological study, Eur J Orthod (2021).

[5] D. Koletsi, A. Madahar, P.S. Fleming, N. Pandis, Statistical testing against baseline was common in dental research, J Clin Epidemiol 68(7) (2015) 776-81.

[6] S. Gratsia, D. Koletsi, P.S. Fleming, N. Pandis, Statistical testing against baseline in orthodontic research: a meta-epidemiologic study, European Journal of Orthodontics 41(2) (2019) 165-171.

[7] N. Pandis, A. Polychronopoulou, P. Madianos, M. Makou, T. Eliades, Reporting of research quality characteristics of studies published in 6 major clinical dental specialty journals, Journal of evidence-based dental practice 11(2) (2011) 75-83.

[8] S.N. Papageorgiou, G.N. Antonoglou, C. Martin, T. Eliades, Methods, transparency and reporting of clinical trials in orthodontics and periodontics, Journal of Orthodontics 46(2) (2019) 101-109.

[9] N. Pandis, Multiplicity 3: repeated measures, Am J Orthod Dentofacial Orthop 143(5) (2013) 747-8.

[10] G. Molenberghs, Verbeke, Geert, Models for Discrete Longitudinal Data, 1 ed., Springer, New York, NY2005.

[11] G. Verbeke, Molenberghs, Geert, Linear Mixed Models for Longitudinal Data, 1 ed., Springer, New York, NY2000.

[12] N. Pandis, P.S. Fleming, C. Katsaros, J.P.A. Ioannidis, Dental Research Waste in Design, Analysis, and Reporting: A Scoping Review, J Dent Res (2020) 22034520962751.

Journal Pre-proof

[13] G.A. Ballinger, Using Generalized Estimating Equations for Longitudinal Data Analysis, Organizational Research Methods 7(2) (2016) 127-150.

[14] K.F. Schulz, D.A. Grimes, Multiplicity in randomised trials I: endpoints and treatments, Lancet 365(9470) (2005) 1591-5.

[15] N. Pandis, Multiplicity 1: subgroup analyses, Am J Orthod Dentofacial Orthop 143(3) (2013) 439-41.

[16] N. Pandis, Multiplicity 2: multiple treatments and multiple outcomes, Am J Orthod Dentofacial Orthop 143(4) (2013) 589-91.

[17] K. Dwan, C. Gamble, P.R. Williamson, J.J. Kirkham, Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review, PloS one 8(7) (2013) e66844.

[18] J.M. Bland, D.G. Altman, Comparisons against baseline within randomised groups are often used and can be highly misleading, Trials 12 (2011) 264.

[19] N. Lu, Y. Han, T. Chen, D.D. Gunzler, Y. Xia, J.Y. Lin, X.M. Tu, Power analysis for cross-sectional and longitudinal study designs, Shanghai Arch Psychiatry 25(4) (2013) 259-62.

[20] J.G. Ibrahim, G. Molenberghs, Missing data methods in longitudinal studies: a review, Test (Madr) 18(1) (2009) 1-43.

[21] A.M. De Livera, S. Zaloumis, J.A. Simpson, Models for the analysis of repeated continuous outcome measures in clinical trials, Respirology 19(2) (2014) 155-161.

[22] S.L. Zeger, K.Y. Liang, An overview of methods for the analysis of longitudinal data, Stat Med 11(14-15) (1992) 1825-39.

[23] G.M. Fitzmaurice, C. Ravichandran, A primer in longitudinal data analysis, Circulation 118(19) (2008) 2005-10.

[24] P.S. Albert, Longitudinal data analysis (repeated measures) in clinical trials, Statistics in Medicine 18(13) (1999) 1707-32.

[25] P. Schober, T.R. Vetter, Repeated Measures Designs and Analysis of Longitudinal Data: If at First You Do Not Succeed-Try, Try Again, Anesthesia & Analgesia 127(2) (2018) 569-575.

[26] L. Hanin, Why statistical inference from clinical trials is likely to generate false and irreproducible results, BMC Medical Research Methodology 17(1) (2017) 127.

[27] N.L. Kerr, HARKing: hypothesizing after the results are known, Personality and Social Psychology Review 2(3) (1998) 196-217.

[28] E. Lempesi, D. Koletsi, P.S. Fleming, N. Pandis, The reporting quality of randomized controlled trials in orthodontics, The journal of evidence-based dental practice 14(2) (2014) 46-52.

Journal Pre-proof

[29] J. Kiriakou, N. Pandis, P. Madianos, A. Polychronopoulou, Assessing the reporting quality in abstracts of randomized controlled trials in leading journals of oral implantology, The journal of evidence-based dental practice 14(1) (2014) 9-15.

[30] N. Pandis, A. Polychronopoulou, T. Eliades, An assessment of quality characteristics of randomised control trials published in dental journals, Journal of dentistry 38(9) (2010) 713-21.

[31] T.R. Vetter, P. Schober, Regression: The Apple Does Not Fall Far From the Tree, Anesthesia & Analgesia 127(1) (2018) 277-283.

[32] T.A. Lang, D.G. Altman, Basic statistical reporting for articles published in biomedical journals: the "Statistical Analyses and Methods in the Published Literature" or the SAMPL Guidelines, International Journal of Nursing Studies 52(1) (2015) 5-9.

Journal Pre-proof

Legends

Table 1 Cross-tabulation between optimal/suboptimal analysis and the study characteristics (* includes 2 classified as unclear **Fisher test)

Table 2 shows the analyses implemented when using the optimal and the suboptimal approach *Mixed models include (Hierarchical, multilevel modelling, GLMs, multilevel linear regression, binomial random effects, logistic mixed effect regression), GEE includes regression with robust variances and RM-ANOVA includes MANOVA and mixed ANOVA

Figure 1 Trial identification flow diagram

Figure 2 A bar plot showing the counts of optimal/suboptimal analysis per specialty journal.

Supplemental Table 1 Search strategy

Supplemental Table 2

Journal Pre-proof

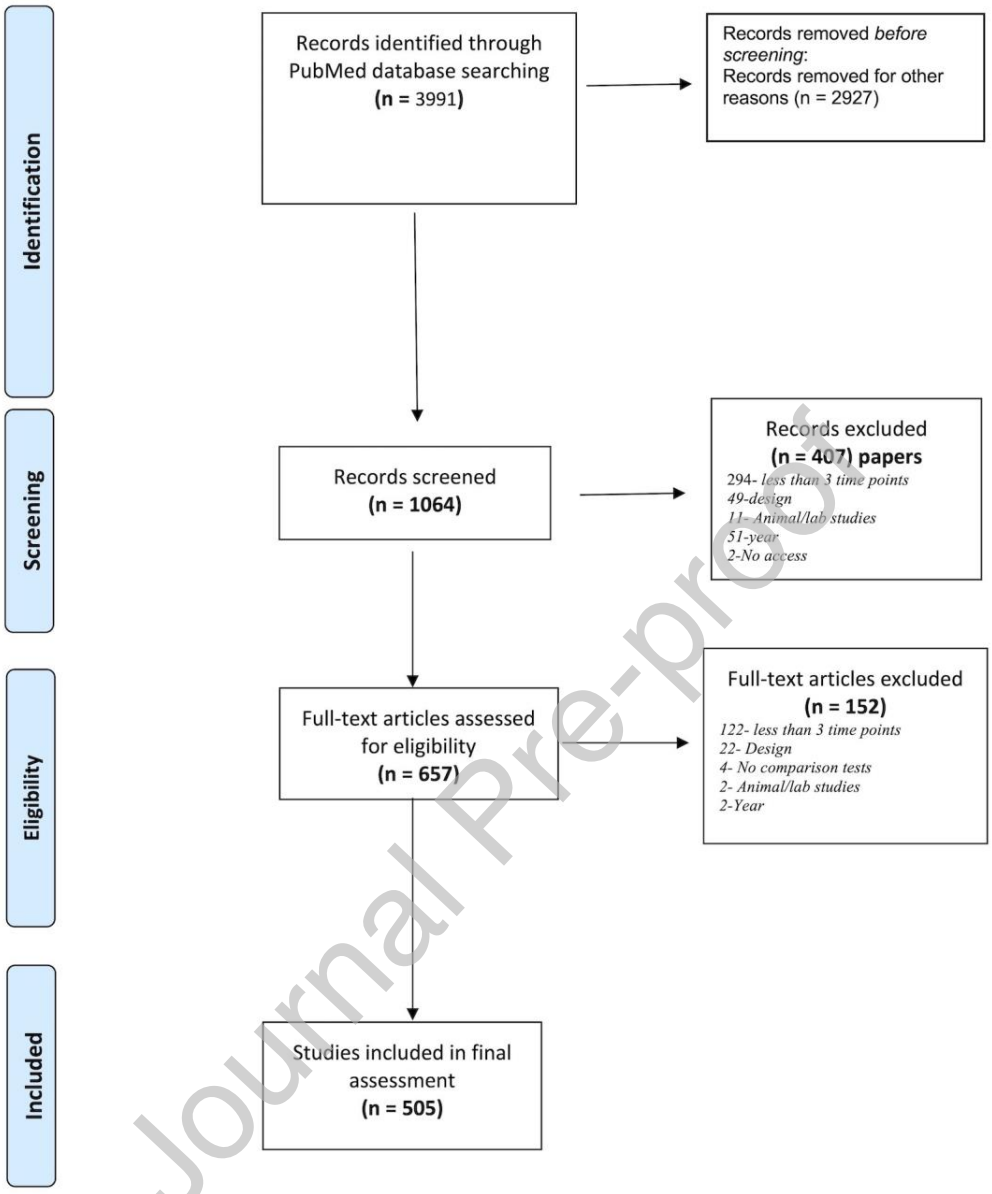


Fig. 1

Journal Pre-proof

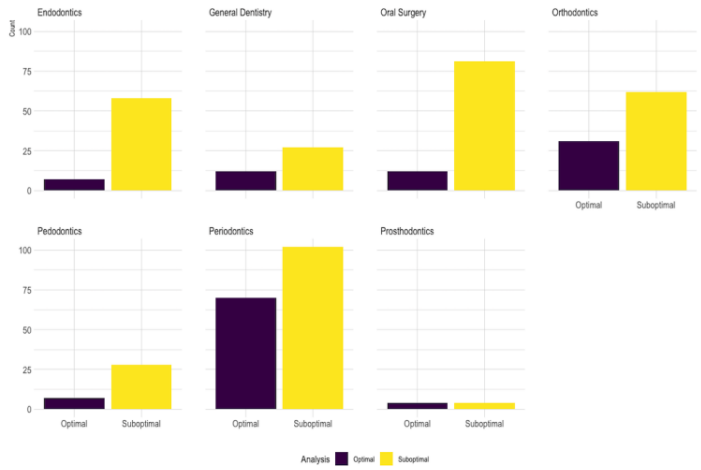


Fig. 2

Journal Pre-proof

Journal Pre-proof

Study characteristics	Optimal analysis						P-value
		No (n= 362 *, 71.7%)		Yes (n= 143, 28.3%)		Total (505) %	
The continent of the first author	Americas	72	64.9%	39	35.1%	(111) 22%	0.02
	Asia and Others	125	79.6%	32	20.4%	(157) 31.1%	
	Europe	165	69.6%	72	30.4%	(237) 46.9%	
Specialty	General Dentistry	27	69.2%	12	30.8%	(39) 7.7%	**<0.001**
	Endodontics	58	89.2%	7	10.77%	(65) 12.9%	
	Orthodontics	62	66.7%	31	33.3%	(93)18.4%	
	Pedodontics	28	80%	7	20%	(35) 6.9%	
	Periodontics	102	59.3%	70	40.7%	(172) 34.1%	
	Prosthodontics	4	50%	4	50%	(8) 1.6%	
	Oral Surgery	81	87.1%	12	12.9%	(93) 18.4%	
Publication year	2016	66	64.7%	36	35.3%	(102) 20.2%	0.46
	2017	82	75.9%	26	24.1%	(108) 21.4%	
	2018	73	73%	27	27%	(100) 19.8%	
	2019	68	73.1%	25	26.9%	(93) 18.4%	
	2020	73	71.6%	29	28.4%	(102) 20.2 %	
Study design	Split Mouth	66	78.6%	18	21. 4%	(84) 16.6%	0.13
	Cross-over	7	53.85%	6	46.15%	(13) 2.6%	
	Parallel	289	70.8%	119	29.2%	(408) 80.8%	
Centre	multi	20	66.7%	10	33.3%	(30) 5.9%	0.55
	single	342	72%	133	28 %	(475) 94.1%	
Number of analyses	1	124	67.4%	60	32.6%	(184) 36.7%	0.46
	2	129	76.3%	40	23.7%	(169) 33.7%	
	3	58	69.9%	25	30.1%	(83) 16.5%	
	≥4	47	71.2%	19	28.8	(66)13.2%	
Statistician Involvement	yes	58	59.2%	40	40.8%	(98) 19.4%	<0.01
	no	304	74.7%	1043	25.3%	(407) 80.6%	
Protocol	yes	167	67.6%	80	32.4	(247) 48.9%	0.05

Journal Pre-proof

registration	no	195	75.6%	63	24.4%	(258) 51.1%	
Funding	yes	183	64.2%	102	35.8%	(285) 56.44%	**<0.001**
	no	179	81.4%	41	18.6%	(220) 43.56%	
Significance of the outcome	Non-significant	144	59.5%	98	40.5%	(242) 39.46%	**<0.001**
	Significant	218	82.9%	45	17.1%	(263) 60.54%	

Table 1

Journal Pre-proof

Table 2

Optimal analysis N= 143 (28.3%)		**Suboptimal analysis N= 359 (71.7%)**		**Unclear analysis N=3 (0.005%)**
Brunner- Langer model	3(0.59)	ANOVA/Other	52 (10.27)	
		Kruskal Wallis	35 (6.9)	
Mixed models	58(11.49)	Fisher/Chi^2	27 (5.15)	
		Mann Whitney/ t-test	202(40)	
Generalized estimating equation (GEE)	11 (2.18)	Paired t-test / Wilcoxon signed-rank test	27 (5.35)	
RM-ANOVA, ANCOVA, Friedman test, SPANOVA	66 (13.27)	multiple linear regression analysis	2(0.4)	
		Median regression	1(0.2)	
log-rank test/other survival analysis	4 (0.61)	multiple randomization tests	2(0.4)	
		polychoricordered logistic regression	2(0.4)	
		Poisson regression	1(0.2)	
		McNemar's test	8 (1.58)	

Journal Pre-proof