ACTA ODONTOLOGICA SCANDINAVICA
2019, VOL. 77, NO. 3, 181–183
https://doi.org/10.1080/00016357.2018.1522448

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Evaluation of dental trials comparing baseline differences using *p* values

Umesh Wadgave[a], Mahesh R. Khairnar[a], Tanvi S. Kadu[b] and Yogesh Wadgave[c]

[a]Department of Public Health Dentistry, Bharati Vidyapeeth Deemed University Dental College and Hospital, Sangli, India; [b]Bharati Vidyapeeth Deemed University Dental College and Hospital, Sangli, India; [c]Department of Community Medicine, Seth G. S. Medical College, Mumbai, India

ABSTRACT

Introduction: Significance testing for comparison of the baseline differences between the intervention arms has received a strong condemnation. The goal of this study was to assess the prevalence of randomized controlled trials (RCTs) comparing the baseline characteristics between intervention groups using significance tests in top ten impact factor dental journals.

Materials and methods: RCTs published in 10 high impact factor dental journals were searched in PubMed database. Literature search was limited to time duration of 5 years from September 2012 to August 2017.

Results: We analysed 521 RCTs after excluding 47 non-RCT articles from the total of 568 articles. Baseline demographic characteristics table was not reported in 45.9% of the RCTs and 26.2% of the RCTs did not report table of baseline clinical characteristics. In 38.9% of the studies, significance testing was employed to compare baseline differences between the intervention arms.

Conclusions: Many trials published in the reputed dental journals failed to follow the recommendations of CONSORT statement regarding reporting of baseline tables and avoiding comparison of baseline differences with significance test.

ARTICLE HISTORY
Received 25 July 2018
Revised 30 August 2018
Accepted 7 September 2018

KEYWORDS
Baseline comparison; biostatistics; dental; randomized controlled trials; test of significance

Introduction

Randomized controlled trials (RCTs) are considered the 'gold standard' of clinical trials and provide most reliable evidence for assessing safety and efficacy of a treatment. In every RCT, randomization ensures chance allocation of the participants to different intervention arms and it is expected that baseline covariates get equally distributed between the different intervention groups [1,2]. However, some important covariates may not always be balanced between intervention groups following randomization [2,3]. Hence, it is necessary to present baseline information on prognostic factors [4]. CONSORT statement recommends a table showing baseline demographic and clinical characteristics for each group. However, it disapproves significance testing of baseline differences by stating that 'tests of baseline differences are not necessarily wrong, just illogical. Such hypothesis testing is superfluous and can mislead investigators and their readers' [5]. Altman stated that 'performing a significance test to compare baseline variables is to assess the probability of something having occurred by chance when we know that it did occur by chance' [6]. Despite criticism for several decades, it is observed that many recent clinical trials compare the baseline differences with significance tests [7–9]. There are no studies found in literature assessing the prevalence of dental clinical trials comparing baseline characteristics between intervention groups using significance tests. The purpose of this study was to assess the prevalence of RCTs comparing the baseline characteristics between intervention groups using significance tests in top ten impact factor dental journals. And also to estimate the proportion of studies reporting baseline demographic and clinical characteristics in a table.

Materials and methods

This study included RCTs from top 10 highest-impact factor dental journals according to the 2016 Journal Citation Reports. PubMed database was searched for RCTs published in selected 10 dental journals. The search strategy used was randomly [All Fields] OR ('random allocation'[MeSH Terms] OR ('random'[All Fields] AND 'allocation'[All Fields]) OR 'random allocation'[All Fields] OR 'randomized'[All Fields])) OR randomized [All Fields]) AND (Clinical Trial[ptyp] limiting to time period of 5 years from September 2012 to August 2017. Following 10 high impact factor dental journal titles were included: Oral Oncology; Journal of Dental Research; Dental Materials: Official Publication of the Academy of Dental Materials; International Journal of Oral Science; Clinical Oral Implants Research; European Journal of Oral Implantology; Journal of Clinical Periodontology; Journal of Dentistry; Journal of Periodontology and International Endodontic Journal. Journal titled Periodontology 2000 was excluded as it publishes only review articles and hence we

CONTACT Umesh Wadgave dr.w.umesh@gmail.com Department of Public Health Dentistry, Bharati Vidyapeeth Deemed University Dental College and Hospital, Wanlesswadi, Sangli 416416, Maharashtra, India

Table 1. Descriptive characteristics of selected studies.

	N	%
Type of study		
Human trials	479	91.9
In vitro studies	26	4.9
In situ studies	14	2.6
Animal studies	2	0.38
Total	521	100
Baseline demography table (Only for human trials)		
Present	259	54.1
Absent	220	45.9
Total	479	100
Baseline clinical table		
Present	384	73.8
Absent	137	26.2
Total	521	100
Baseline significance testing		
Done	203	38.9
Not done	318	61.03
Total	521	100

Table 2. Year wise comparison of studies reporting baseline tables and significance testing.

		Baseline findings		
Year	RCTs N	Demography table (in human trials)[a] N (%)	Clinical table N (%)	Significance testing N (%)
2012	45	19 (51.4)	33 (73.3)	19 (42.2)
2013	175	83 (54.2)	124 (70.9)	66 (37.7)
2014	130	63 (49.6)	101 (77.7)	56 (43.1)
2015	119	60 (53.6)	86 (72.3)	43 (36.1)
2016	39	25 (67.6)	30 (76.9)	13 (33.3)
2017	13	9 (69.2)	10 (76)	4 (30.3)
Total	521	259	384	203

[a]Baseline demographic tables are necessary to be mentioned only in human trials.

included 11th ranked journal title International Endodontic Journal.

Our search strategy in PubMed database resulted in 568 articles totally. Full texts of all the selected articles were retrieved and the two authors (UW and MK) independently screened all the articles against eligibility criteria. Both the authors analysed full text, and any disagreements were resolved through discussion and came to a consensus. Data from the selected studies were independently extracted by two authors (UW and MK). Significance testing of baseline characteristics was evaluated in both baseline table (demographic and clinical) and in the text. Data analysis was performed using SPSS software version 16 (SPSS, Chicago, IL). Frequency and percentage distributions were used to estimate the proportion of articles reporting baseline demographic or clinical characteristics of the study subjects, and reporting use of significance tests to compare these basic characteristics between the study groups.

Results

Totally 568 articles were retrieved from PubMed database. Forty-seven articles were excluded because they were non-RCTs. Remaining 521 articles were analysed.

Majority of the RCTs included for the study were human trials (91.9%). Among the human trials, about 45.9% studies failed to show the table with baseline demographic characteristics. However, 26.2% studies failed to show the table with baseline clinical characteristics. This result shows that more studies reported baseline tables of clinical characteristics (73.8%) than demographic characteristics (54.07%). Totally, 38.9% of studies did the significance testing for comparison of baseline covariates/characteristics between the intervention arms (Table 1).

In the year 2016, the highest percentage (67.6%) of the human trials reported baseline demographic table; whereas minimum percentage was observed in the year 2014 (49.6%). Year-wise comparison also revealed that every year about 70–80% of studies reported table of baseline clinical characteristics. More number of studies applied significant tests to compare the baseline differences in the year 2014 (43.1%) followed by 2012 (42.2%) (Table 2).

Discussion

Explicit and error-free reporting of patients' characteristics in RCTs is crucial for generalizing the results [10]. CONSORT statement recommends a table showing baseline demographic and clinical characteristics for each group. Such information allows readers to judge how relevant the results of a trial might be to an individual patient. In this study, 45.9% of the human trials failed to report the baseline demographic table and 26.2% trials failed to report baseline clinical characteristics table of each group. Altman and Doré and Assmann et al. found that 8% trials failed to report the baseline table in reputed medical journals (Annals of Internal Medicine, NEJM, The Lancet, JAMA and BMJ) [11,12]. Austin et al. showed that 3.5% RCTs published in the NEJM, JAMA, Lancet and BMJ in 2007 did not report the baseline table [1]. Knol et al. reviewed RCTs between 2008 and 2010 in seven medical and cardiovascular journals (Annals of Internal Medicine, BMJ, Circulation, European Heart Journal, JAMA, Lancet and NEJM) and reported that 5% trials did not have baseline table. Recently in 2017, Peterson et al. found that 20% trials of 10 high-impact sports medicine journals did not report the baseline table [7]. Findings from the above studies clearly indicate that dental journals less frequently reported the baseline characteristics table as compared to medical journals.

In this study, 38.6% of the trials performed significance testing to compare the baseline differences. In previous studies, 64.8% trials of ten high impact sports medical journal of 2005 and 2015 [7], 38% trials of seven major general medical and cardiovascular journals between 2008 and 2010 [8], 38.2% RCTs published in four major medical journals (NEMJ, JAMA, Lancet and BMJ) in 2007 [1], 50% trials of five major medical journals (Annals of Internal Medicine, NEJM, The Lancet, JAMA and BMJ) during 1997 [12] and 58% of four major medical journals in 1987 [11] reported significance testing of baseline differences between the groups.

Researchers substantiate the use of significance tests for comparing the baseline differences in RCTs with two reasons: first, to test the success of randomization and; second, to assess whether the baseline differences are real or important

which can influence (confound) the study results [9]. Contrastingly other scholars consider the application of significance test after randomization as unnecessary and have given justification for the above two reasons [6,13]. First, there is no clear threshold of the baseline differences to test the success of randomization. Randomization is to avoid selection bias and even a successful randomization shows statistically significant differences at baseline. Therefore, the presence of statistically significant difference does not indicate the success of randomization and its absence will not compensate for a failed randomization process [13,14]. Second, randomization does not guarantee the baseline balance of covariates and any differences observed in baseline characteristics are because of chance. Comparing baseline differences post randomization with tests of significance will disregard the chance allocation of participants [6,9]. CONSORT statement has also opposed such comparison stating it as illogical and misleading [5]. Evaluation of statistical significances should be applied only for the main research questions of the clinical trial, i.e. to compare the primary and secondary outcome variables between the treatment groups. Baseline characteristics are not the primary outcome variables.

Comparable intervention arms are important at baseline and should be presented when studies are reported. It is advised to look for the meaningful difference at baseline rather than statistically significant differences because trials are not powered for this purpose [9,15,16]. Meaningful differences are the baseline differences that can influence (confound) the results of a trial [9]. Meaningful differences can be identified carefully with sound clinical knowledge, research evidence and common sense [6,14]. They should be adjusted using analysis of covariance regardless of the presence or absence, or significance or non-significance, of baseline imbalances [6,12]. One should also be aware of inappropriate selection of covariates for adjustment can result in erroneous estimates [1].

Conclusion

It has been found that several RCTs in the leading dental journals have failed to report baseline tables and allowed erroneous statistics of using significance testing to compare baseline characteristics between intervention arms. Our study suggests greater consistency in reporting of RCTs and researchers should make the statistical concepts clear before considering them.

Disclosure statement

All the authors declare no conflict of interest.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

ORCID

Umesh Wadgave http://orcid.org/0000-0001-5359-834X

References

[1] Austin PC, Manca A, Zwarenstein M, et al. A substantial and confusing variation exists in handling of baseline covariates in randomized controlled trials: a review of trials published in leading medical journals. J Clin Epidemiol. 2010;63:142–153.

[2] Abel U, Koch A. The role of randomization in clinical studies: myths and beliefs. J Clin Epidemiol. 1999;52:487–497.

[3] Senn S. Testing for baseline balance in clinical trials. Stat Med. 1994;13:1715–1726.

[4] Raab GM, Day S, Sales J. How to select covariates to include in the analysis of a clinical trial. Control Clin Trials. 2000; 21: 330–342.

[5] Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg. 2012;10:28–55.

[6] Altman DG. Comparability of randomised groups. Stat Health. 1985;34:125. [cited 2017 Nov 12]Available from: http://www.jstor.org/stable/10.2307/2987510?origin=crossref.

[7] Peterson RL, Tran M, Koffel J, et al. Statistical testing of baseline differences in sports medicine RCTs: a systematic evaluation. BMJ Open Sport Exerc Med. 2017;3:e000228. Internet]. [cited 2017 Nov 11]Available from: http://www.ncbi.nlm.nih.gov/pubmed/28761712.

[8] Knol MJ, Groenwold RHH, Grobbee DE. P-values in baseline tables of randomised controlled trials are inappropriate but still common in high impact journals. Eur J Prev Cardiol. 2012;19:231–232. [Internet]. [cited 2017 Nov 11]Available from: http://www.ncbi.nlm.nih.gov/pubmed/22512015.

[9] de Boer MR, Waterlander WE, Kuijper L, et al. Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. Int J Behav Nutr Phys Act. 2015;12:4. Available from: http://www.ijbnpa.org/content/12/1/4.

[10] Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?" Lancet. 2005;365:82–93.

[11] Altman DG, Doré CJ. Randomisation and baseline comparisons in clinical trials. Lancet. 1990;335:149–153.

[12] Assmann SF, Pocock SJ, Enos LE, et al. Subgroup analysis and other (mis)uses of baseline data in clinical trials. Lancet. 2000;355: 1064–1069. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0140673600020390.

[13] Senn S. Seven myths of randomisation in clinical trials. Stat Med. 2013;32:1439–1450.

[14] Fives A, Russell DW, Kearns N, et al. The Role of Random Allocation in Randomized Controlled Trials: Distinguishing Selection Bias from Baseline Imbalance. J MultiDiscip Eval. 2013;9: 33–42.

[15] Lavori PW, Louis TA, Bailar JC, et al. Designs for experiments — parallel comparisons of treatment. N Engl J Med. 1983;309: 1291–1299. Available from: http://www.ncbi.nlm.nih.gov/pubmed/6633587%5Cn http://www.nejm.org/doi/abs/10.1056/NEJM198311243092105.

[16] Ciolino JD, Martin RH, Zhao W, et al. Measuring continuous baseline covariate imbalances in clinical trial data. Stat Methods Med Res. 2015;24:255–272. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid =4280338&tool=pmcentrez&rendertype=abstract%5Cn http://smm.sagepub.com/cgi/doi/10.1177/0962280211416038.