Guidelines for standardizing and increasing the transparency in the reporting of biomedical research

Amir Maroof Khan

Department of Community Medicine, University College of Medical Sciences, Delhi, India
Correspondence to: Amir Maroof Khan, Associate Professor. Department of Community Medicine, University College of Medical Sciences, Delhi, India. Email: khanamirmaroof@yahoo.com.

Abstract: There is a lack of awareness about the guidelines for standardized and transparent reporting of biomedical research, among the medical professionals. This paper aims to familiarize the clinical researchers and practitioners regarding the issues related to transparency and the evolving guidelines for standardizing them, in the reporting of biomedical research. A narrative review method is adopted here, primarily based on the EQUATOR and SAMPL guidelines for reporting studies and statistical analyses. As study methods and statistical approaches support each other, their reporting practices as per the standardized guidelines have been dealt here in a congruous manner.

Keywords: Biomedical research; biostatistics; medical writing; guideline

Introduction

"Because what is ultimately at stake is public trust in science" (1).

Scientific scrutiny of the 'evidence' in medical research, in order to improve its scientific reliability, show that most of the medical research findings and their interpretations are false and lead to colossal waste of research resources (2,3). Among all the research practices suggested to improve the reliability of medical research, improving and standardizing reporting is one of the key elements (4).

Efforts have been made to standardize reporting of the studies conducted, in user-friendly checklist formats, to increase reliability in reporting research. These guidelines have now been adopted by many medical journals and funding organizations, resulting in increased transparency of clinical study reporting in journals that have adopted them (5). Thus, it is imperative for clinical researchers to be aware about them, their significance in medical research, and use it when and where applicable.

The current article is based primarily on these guidelines to inform and empower the clinical researchers about biomedical research reporting.

Guidelines to standardize and increase transparency in reporting biomedical literature

When it comes to reporting biomedical research, the statistical analyses reporting and study design details reporting go hand in hand.

The "Enhancing the Quality And Transparency of health Research" (EQUATOR) network has published more than 350 guidelines for Health Research Reporting (6). With respect to generic formats of study designs, the common ones are the "Consolidated Standards for Reporting Trials" (CONSORT) for randomized trials (5), "Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)" for observational studies (7), "Standards for Reporting of Diagnostic" (STARD) for diagnostic studies (8), "Preferred Reporting Items for Systematic Reviews and Meta-Analyses" (PRISMA) for systematic reviews and meta-analyses (9), and "Standard Protocol Items: Recommendations for Interventional Trials" (SPIRIT) for defining protocol items for clinical trials (10). More and more specialised guidelines have been developed such as CONSORT non-inferiority for reporting of non-inferiority and equivalence trials (11), the RECORD

statement for the reporting of studies conducted using observational routinely collected health data (12), and so on. Apart from these, there are other guidelines too that deserve a mention such as "Statistical Analyses and Methods in the Published Literature" or SAMPL guidelines for reporting statistical analyses (13), and "Strengthening Analytical Thinking for Observational Studies" (STRATOS) for reporting observational studies (14). Readers are referred to the EQUATOR network website for accessing the specific reporting guidelines for their specific research designs.

The appropriate reporting guideline/s should be referred to at the planning stage of the medical research itself. Prior to the initiation of the study, this will guide the researcher to better consider all the important elements to be included in the study design (10).

Here we describe and discuss, from the above mentioned guidelines, some common elements to be reported to enhance transparency in biomedical research.

Randomization

It has been reported that around one third of the Randomized Controlled Trials (RCT) did not report the method of randomization (15). Selection of study subjects, or allocating study subjects to a particular group, is an important step to eliminate 'selection bias' in any biomedical research. It is important to identify and report the target population and then the sampling or random allocation technique. On the other hand, the term randomization should not be used loosely. It should be mentioned, only if some randomization technique, such as computer generated random numbers, is used.

Blinding

Blinding and allocation concealment are different entities. Both are essential to eliminate selection biases (16). Blinding is when the participant and/or the investigator does not know about the treatment being administered. Allocation concealment is different from blinding. This technique e.g., sealed envelope method, doesn't allow the investigator which group the next patient will belong to.

Sample-size estimation

A study reported that 43% of the published RCT did not report all the parameters required to calculate the sample size (17). A study with inadequate sample size may not

be able to detect a difference, even if there is a difference between them, whereas a study with a larger than adequate sample size may detect a difference, even when it does not exist. Appropriate sample-size should be computed based on the type of study and reported a priori.

As sample sizes are calculated based on certain estimates and assumptions, it is important to report the estimates used, along with the source of this data. These estimates may be obtained from published studies, or from unpublished data available from the medical records department, or a pilot study can be done to generate some data to arrive at an estimate. Feasibility aspects and sample size computation using certain mathematical estimates go hand in hand to achieve a balance in arriving at the final sample size. As different methods are available to compute sample sizes, it is necessary that the sample size calculation formula, or any software which has been used, should be reported.

Eligibility criteria for participants

The inclusion and exclusion criteria should be categorically reported with the operational definitions for each of these criteria. Most clinical studies compare two groups of participants, based on the outcome of interest, or on the exposures to risk factors or on the status of the treatment received. It is important to apply the same eligibility criteria to both the groups to ensure comparability between them. Once enrolled in the study, its outcomes, benefits and harms both, should be recorded and reported.

The setting and the locations where the participants were enrolled should be reported. The sources and methods of participant selection should be mentioned. Often, matching is done for observational, comparative studies, to control for confounding variables. For such matched case-control or cohort studies the matching criteria should be mentioned.

Statistical analysis

The SAMPL guidelines categorize the statistical analyses common to most of the study types into three sections i.e., preliminary analyses, primary analysis, and supplementary analyses (13).

Preliminary analyses

This includes the statistical processes used to transform or

modify data. This might be required in cases where data is not normally distributed but for the purpose of statistical analyses, normal data is required. How the variables will be used in the study should be identified prior to the data collection. This means whether the variable will be considered independently or whether some other index or indicator will be derived by using more than one variable. Even the categories to be formed out of continuous data, or by collapsing some categories into newer variables, should be defined a priori, preferably by giving some justification for it.

Primary analyses

This includes the following:
(I) Describing the appropriateness of the statistical test used.

Each variable should be summarized using descriptive statistics. In case of proportions, absolute values must be mentioned. Appropriate measures of central tendency (Means for normal distribution and Median for non-normal distribution) and dispersion (Standard deviation for normal distribution and Interquartile range for non-normal distribution) should be reported. When the variation in the sample is to be depicted, standard deviation should be used and not standard error. Standard error is used when population estimates are desired based on the sample estimates.

Clearly report the variables for which the specific statistical tests were used. Just giving a list of statistical tests without specifying the variables on which these will be used or just mentioning that 'appropriate statistical tests will be/were used for analysis' is incomplete reporting.

Care should be taken to check the type of distribution for continuous variables, and when non-normal or skewed distribution is detected, non-parametric tests should be applied. Similarly, paired data, such as in before-after studies or matched pair studies, need to be analysed using paired tests.

(II) Multiple comparisons and analyses.

If the research objective is to have multiple group (more than two) comparisons, then appropriate statistical tests should be used. Multiple two-group comparison tests cannot be used here. There are post-hoc tests i.e., after the

multiple group comparison statistical test has been applied, for two-group comparisons. This is because certain adjustments are needed to do two group comparison testing in such cases. Hence it is important to identify at the designing stage, how many group comparisons are required by the primary objective of the study and decide the statistical test appropriately.

(III) Name the statistical test package used.

As there are different approaches in statistics, even for the same statistical task, there might be certain differences in the various statistical packages. In the absence of this information, the same statistical results cannot be reproduced. To maintain reliability of the results given by you, it is essential that the statistical software package should be mentioned.

(IV) Outlying data.

How to treat extreme values is a matter of debate. Should they be included in analysis or should it be removed from analysis. It is better to give the readers both the results, one obtained by including them in the analysis and another by removing them from the analysis. You can present your own arguments towards both the findings, and the reader will also be free to have her own interpretations based on the results presented. Even if any data point is to be excluded, it should be decided a priori, at the time of designing the study and writing the protocol, and not later, as doing it later will make your findings questionable (18).

(V) The hypothesis to be tested should be mentioned a priori.

This is necessary as this guide towards the whole planning and conduct of the study. From the sample size calculation, to the consideration of confounding variables, to deciding about the statistical analysis, everything in the study rests on it. All other outcomes from the study, are incidental or chance findings. The study was not designed for these secondary outcomes, with the rigour that it was designed and implemented for the primary outcome. These post-hoc analyses can be used to explore further relationships between various variables. What was decided a priori and what was done as post-hoc analyses should be reported to maintain full transparency of the research outputs (18).

Supplementary analyses

Sensitivity analysis

Statistics is based on certain assumptions: about confounding, about comparability, about distribution, about population, and different statistical approaches use different assumptions to give the statistical output. We do not know how robust the statistical output that we have obtained using a particular statistical approach is unless we modify these assumptions and see whether the statistical output remains within certain acceptable limits or changes drastically. There are analytical techniques which can be used to test for sensitivity testing. This can be carried out for observational as well as interventional studies and are recommended by the medical research reporting guidelines.

Subgroup analysis and Post-hoc analysis should also be reported, as has been discussed above.

Missing data

Non-response rates should be reported. If possible, their background characteristics should be noted so as to be able to draw valid inferences by assessing as to what extent were they different from the ones included in the study.

In follow-up studies, attrition or dropouts is a challenge. There are various methods of imputation for handling missing data. The missing data should be described and discussed as well as the methods employed to handle it should be reported (19).

It would be better if some steps to minimize the non-response rates and dropouts are decided a priori and implemented during the conduct of the study.

Model building

Many medical research studies now develop and report multivariate regression models to adjust for the confounding variables. The criteria that will be used to include/exclude variables from the model and the model building method such as forward selection or others, must be reported. All regression models, are based on assumptions. It is important to report how the assumptions were tested and whether those assumptions were met or not, when reporting the findings. Report the regression coefficients (beta weights) of each explanatory variable and the associated confidence intervals and P values, preferably in a table. Provide a measure of the model's "goodness-of-fit" to the data (the coefficient of determination, r^2, for simple regression and the coefficient of multiple determination, R^2, for multiple regression).

Planned interim data analysis

Often, researchers, as the study progresses do interim data analysis and see the results. If it is in favour of their hypothesis, they go ahead for reporting or publishing it. This is a wrong approach and leads to false positive findings (2). Any interim analysis, if to be done should be planned a priori, based on some logical argument and not based on intuition. There is no scope for unplanned interim analysis in medical research. Even if it is the case, then it should be clearly reported to maintain transparency of the findings.

Clustering effect

Most of the statistical analyses assume that the units of analysis are independent observations. Let us consider that there are ten blood pressure measurement readings belonging to a single individual at different points of time. These readings cannot be considered as independent. This is the clustering effect, which should be considered and reported, and appropriate statistical tests should be applied to adjust for its effects (19).

Clinical versus statistical significance

P value tells us about the statistical significance, whereas effect size reveals the clinical significance. A smaller P value does not mean that there is a high correlation or strong association. Similarly a study with a large sample size may find a small effect size to be statistically significant. Incorrect interpretation and reporting of P values and giving too much emphasis on P values less than 0.05 when the effect size is clinically meaningless, should be avoided. And effect sizes should be mentioned along with their 95% confidence intervals.

The magnitude of the effect can be assessed by the difference between means, or the difference in proportions, or the risks and risk ratios, the correlation coefficients, and other such measures. For primary outcomes, confidence intervals are preferred over the P values. Even if P values are given, it should be reported as exact values and up to two decimal places. (e.g., P value =0.03) and not as NS (non-significant) or as < or >0.05.

Misinterpretation of the results

Researchers frequently interpret odds ratios (OR) as risk ratios (RR) and report it as such. OR and RR are two different entities. OR overestimates RR. OR are reported

in case-control studies and logistic regression models, while RR in cohort studies.

Reporting Likelihood ratios without reporting the sensitivity and specificity of a test is incomplete reporting and should be avoided.

Point estimates along with standard error should be correctly mentioned for depicting 95% CI (confidence interval). At times, researchers mention point estimate ±SE (standard error) and report it as 95% CI, but this represents 68% CI. A point estimate ±2SE will represent 95% CI.

Existing challenges in standardization and increasing the transparency in the reporting of biomedical research

One of the challenges highlighted is that if the guidelines become too rigid on the reporting standards, it may restrict research (20). It may be argued that an incorrectly or incompletely reported research will harm the spirit of science more than it will benefit it and so these restrictions might be good for advancement of good scientific practices.

Implementation of these guidelines is also challenging. All stake holders including the editors, reviewers, researchers should promote and support the implementation of these guidelines (21).

With regards to increasing transparency, there is a need for promoting data sharing by the medical researchers. This will help in identifying errors in original analyses and to conduct additional analysis (22). Currently, the guidelines do not mention this aspect and there is no general consensus among the researchers on this issue.

Conclusions

Comprehensive and transparent reporting makes research credible, reproducible and help reduce research wastage. Standardized guidelines give directions towards achieving this objective. More and more specialized biomedical research reporting guidelines are being developed which can serve the varied interests of the medical researchers to make research transparent. There is a need to generate interest among the medical researchers and make them aware towards using these guidelines.

Acknowledgements

None.

Footnote

References

1. Enhancing reproducibility. Nat Meth. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved, 2013;10:367.
2. Ioannidis JP. Why most published research findings are false. PLoS Med 2005;2:e124.
3. Macleod MR, Michie S, Roberts I, et al. Biomedical research: increasing value, reducing waste. Lancet 2014;383:101-4.
4. Ioannidis JP. How to make more published research true. PLoS Med 2014;11:e1001747.
5. Schulz KF, Altman DG, Moher D, et al. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. PLoS Med 2010;7:e1000251.
6. The EQUATOR Network. Reporting guidelines. [Internet]. [cited 2017 May 25]. Available online: https://www.equator-network.org/reporting-guidelines/
7. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370:1453-7.
8. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015;351:h5527.
9. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097.
10. Chan AW, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013;346:e7586.
11. Piaggio G, Elbourne DR, Altman DG, et al. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. JAMA 2006;295:1152-60.
12. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med 2015;12:e1001885.
13. Lang TA, Altman DG. Basic statistical reporting for articles published in biomedical journals: the "Statistical Analyses and Methods in the Published Literature" or the SAMPL Guidelines. Int J Nurs Stud 2015;52:5-9.

14. Sauerbrei W, Abrahamowicz M, Altman DG, et al. STRengthening analytical thinking for observational studies: the STRATOS initiative. Stat Med 2014;33:5413-32.

15. Kahan BC, Rehal S, Cro S. Risk of selection bias in randomised trials. Trials 2015;16:405.

16. Nüesch E, Reichenbach S, Trelle S, et al. The importance of allocation concealment and patient blinding in osteoarthritis trials: a meta-epidemiologic study. Arthritis Rheum 2009;61:1633-41.

17. Charles P, Giraudeau B, Dechartres A, et al. Reporting of sample size calculation in randomised controlled trials: review. BMJ 2009;338:b1732.

18. Thiese MS. Observational and interventional study design types; an overview. Biochem Med (Zagreb) 2014;24:199-210.

19. Fernandes-Taylor S, Hyun JK, Reeder RN, et al. Common statistical and research design problems in manuscripts submitted to high-impact medical journals. BMC Res Notes 2011;4:304.

20. Morris R. Reporting standards: Rigid guidelines may restrict research. Nature Publishing Group, a division of Macmillan Publishers Limited, 2012;492:192.

21. Baker D, Lidster K, Sottomayor A, et al. Reproducibility: Research-reporting standards fall short. Nature. Nature Publishing Group, a division of Macmillan Publishers Limited, 2012;492:41.

22. West R. Data and statistical commands should be routinely disclosed in order to promote greater transparency and accountability in clinical and behavioral research. J Clin Epidemiol 2016;70:254-5.