Title: International Multi-stakeholder Consensus Statement on Clinical Trial Integrity

Authors: Cairo Consensus Group on Research Integrity (Authorship details in Table 1)

***Corresponding email:** profkkhan@gmail.com

ABSTRACT

Background: Science integrity initiatives require specific recommendations for randomised clinical trials (RCT).

Objective: To prepare a set of statements for RCT integrity through an international multi-stakeholder consensus.

Methods: Following prospective registration (https://osf.io/bhncy, https://osf.io/3ursn), the consensus was developed via: multi-country multidisciplinary stakeholder group composition and engagement; evidence synthesis of 55 systematic reviews concerning RCT integrity; anonymised two-round modified Delphi survey with consensus threshold based on the average percent of majority opinions; and, a final consensus development meeting.

Results: There were 30 stakeholders representing 14 countries from 5 continents including trialists, ethicists, methodologists, statisticians, consumer representative, industry representative, systematic reviewers, funding body panel members, regulatory experts, authors, journal editors, peer-reviewers and advisors for resolving integrity concerns. Delphi survey response rate was 86.7% (26/30 stakeholders). There were 111 statements (73 stakeholder-provided, 46 systematic review-generated, 8 supported by both) in the initial long list, with 8 additional statements provided during the consensus rounds. Through consensus the final set consolidated 81 statements (49 stakeholder-provided, 41 systematic review-generated, 9 supported by both). The entire RCT life cycle was covered by the set of statements including general aspects (n=6), design and approval (n=11), conduct and monitoring (n=19), reporting of

protocols and findings (n=20), post-publication concerns (n=12), and future research and development (n=13).

Conclusion: Implementation of this multi-stakeholder consensus statement is expected to enhance RCT integrity.

Tweetable abstract: An international clinical trial integrity statement, developed through multi-stakeholder consensus, needs implementation.

INTRODUCTION

The essence of the multiple concepts and terms related to research integrity[1–5] boils down to responsible research conduct through compliance with ethics and professional standards.[1,6] A working definition of science integrity clarifies the crucial role of "ensuring transparency at all stages of design, execution, and reporting".[3] Existing integrity initiatives[7–9] provide general statements about how to promote responsible research conduct.

In health effectiveness research, as randomised clinical trials (RCTs) and their systematic reviews are at the highest level of the evidence validity hierarchy, preserving RCT integrity is a priority.[10–12] The high rates of questionable research practices in integrity surveys,[12,13] and the growing number of allegations of data fabrication in retractions[14] have shaken practitioner and public confidence. Not all such cases are due to deliberate misconduct.[15] RCT integrity, however, is under threat from a mix of unintentional errors, faulty methodology, lack of awareness of research ethics, poor writing skills, pressure to publish, etc.[1,11,16–18] To our knowledge, apart from the International Council on Harmonisation of technical requirements for registration of pharmaceuticals,[19] the research integrity initiatives[7–9] are not specific to RCTs. This makes it difficult for the clinical academic institutions, research funding bodies, and publishing organisations to target RCTs for improving their integrity standards. Thus, there is an urgent need for RCT community alignment in this area.[20]

To address the need for an updated and specific set of integrity statements relating to responsible research conduct for RCTs, we undertook an international multi-

stakeholder consensus focussing on the transparency required at the various stages of their planning, execution and reporting.

METHODS

Following prospective registration (https://osf.io/bhncy), we developed this international consensus statement on RCT integrity, according to recommended methods,[21–25] using a multi-step approach: a) multi-country multidisciplinary stakeholder group composition and engagement; b) evidence synthesis of systematic reviews of RCT integrity; c) anonymised two-round modified Delphi survey; and, d) a final consensus development meeting.

a. Establishment of the international multi-stakeholder group

In August 2021, six months ahead of the proposed consensus meeting, an international stakeholder group was carefully composed selecting members based on their knowledge and experience to encompass all the critical aspects of the RCT research lifecycle. A clinical trial was defined as a study design that randomly assigns human participants to one or more interventions and follows them up for critical outcomes to determine the effect of the interventions.[10] Stakeholders were representatives from: relevant professional societies; allied health professions; patient, public and consumer representatives; trialists, statisticians, and methodologists; members and reviewers of ethics, data monitoring and funding committees; peer-reviewers and biomedical journal editors. They were contacted via direct email (see the list of stakeholders and their roles in Table 1). We ensured that none of the participants had any RCT papers subjected to an active expression of concern nor retraction. All stakeholders explicitly

declared their conflicts of interests using the International Committee of Medical Journal Editors (ICMJE) uniform disclosure form (Appendix 1). One non-voting member (DM) was invited to the group for advising on consensus methods and language. Two members of the group were selected as co-convenors (KSK and YK), charged with the responsibility to ensure that all participants developed ownership of the consensus scope and content, engaging them in discussions, constructive debates and resolution of disagreements. Following acceptance of the invitation, online or phone interviews were held with the stakeholders to inform them about the project objectives, and to ask them for their input to the integrity statements.

b. Umbrella review for generating evidence-based statements

For the creation of the initial long list of statements, we conducted a review of systematic reviews on RCT research integrity. The prospectively registered umbrella review (https://osf.io/3ursn) was carried out with a comprehensive search strategy covering major electronic databases (PubMed, Scopus, Cochrane Library and Google scholar) from inception to November 2021 to capture peer-reviewed and grey literature. The review's search and selecting strategy, data extraction, methods for assessing methodological quality, and synthesis of findings have been reported in an accompanying paper.[20] Building on the collated findings, a core group of four stakeholders (AB, PC, MF and KSK) drafted clear, precise, and actionable statements. The statement drafting process was piloted using seven included reviews initially. The deliberations at this stage helped to clarify the distinction between review findings and the resulting statements. Each member of the core stakeholder group first

independently drafted statements, aiming for one action or recommendation per statement, and then finalised them through discussion.

c. Modified Delphi survey

The statements provided by stakeholders were added to those generated from the umbrella review without editing. Together they created the long list for the modified Delphi consensus survey among 30 stakeholders with voting rights deploying a web-based survey tool (www.surveymonkey.com). A seven-point scale was provided to assess the level of agreement with the content of each statement. The scale was anchored between "strongly agree" and "strongly disagree", with "agree", "somewhat agree", "neither agree nor disagree", "somewhat disagree", and "disagree" included as the scaled options for responses. The same scale was used in both survey rounds administered on 30th January and 9th February 2022. The sum of the "strongly agree" and "agree" responses were used to compute an agreement rate for the approval of each individual statement. The responses of the individual stakeholders were kept anonymous throughout the whole process.

We used an objective method to determine the threshold or cut-off for approval of the statements, average percent of majority opinions (APMO).[25] For this computation, a statement was considered as agreed if the majority (>50%) of stakeholders responded "strongly agree" or "agree" on the seven-point scale. A statement was considered as disagreed if the majority (>50%) of stakeholders responded "disagree" or "strongly disagree" on the seven-point scale. The AMPO consensus threshold was calculated as:

sum of majority agreed and majority disagreed statements / total number of responses received x 100%. Statements above the APMO threshold were considered as having reached consensus. For individual statements that reached consensus in each round we computed the strength of the agreement among stakeholders using the interquartile range (IQR).[24] IQR was the difference between first and third quartiles of the stakeholders´ responses on the seven-point scale. It was interpreted as follows: IQR 0 (>50% stakeholders gave the same responses) indicated very good strength of agreement; IQR 1 (>50% stakeholders´ range of responses was ≤2 points of the scale) indicated good strength of agreement; IQR ≥2 (>50% stakeholders´ range of responses was >2 points of the scale) indicated poor strength of agreement. As a sensitivity analysis, we used an arbitrary approval threshold of 70%. Results were analysed using Stata v16 software (StataCorp. 2019, College Station, TX: StataCorp LLC).

Statements not having reached consensus in the first round using the APMO threshold were merged with new statements provided by stakeholders and subjected to the second round of the modified Delphi survey. The statements deemed to have failed to reach consensus because of lack of clarity in language had their wording improved. The statements containing similar information were merged to avoid duplication. First-round agreement rate was provided in the second survey round along with the references to the reviews supporting the statements generated via evidence synthesis. The minor rewording, statement merger and statistical approach in the second round was the same as that used in the first round. The statements that failed to reach consensus were taken for voting to the final consensus development meeting.

To consolidate the provisional statement set, a core group of stakeholders (AB, KSK, MNN, PC, MF) evaluated the statements that had reached consensus for exact or inexact duplications and clarity of meaning. Where the duplication was virtually exact, a single statement was created, making only minor wording changes to clarify or enhance the intended meaning. No major wording changes were introduced to any of the statements that had met the consensus threshold. The statements without consensus were revised in the same manner with a view to improving the clarity of their meaning and to assist in subsequent voting. Thus, an original statement may have been subjected to minor rewording or merger with other statements various times through the different consensus rounds. The list of statements resulting from the above process, both those having reached consensus and those not having done so, was tabulated and circulated to all the participants with the agreement ratings and the underpinning references to reviews for the consensus development meeting.

d. Consensus development meeting

All stakeholders were invited to the meeting, which was attended by 24 participants in person, 6 participants virtually for the entire day, and DM in person as an advisor. The provisional statement set tabulated above was shared with the participants together with an initial draft of this manuscript. At the meeting, held in Cairo, Egypt, on the 22nd February 2022, statements that were classified as not having reached consensus in the two-round Delphi survey were individually discussed. Stakeholders decided on the agreement rate to be used as the threshold for exclusion and voted anonymously using

an electronic system (Zoom meeting software) to select statements for the final set. The breakdown of statements into the various stages of the RCT research lifecycle was agreed with the stakeholder group. This included subheadings general, design and approval, conduct and monitoring, reporting of protocols and findings, post-publication concerns, and future research and development. In tabulation of the final set, the strength of evidence assessed via a modified AMSTAR-2 score[26] was provided for the statements underpinned by systematic reviews.

PATIENT AND PUBLIC INVOLVEMENT

One patient representative was a stakeholder in the consensus group to provide input as a trial participant. Three stakeholders had prior experience in patient, public and consumer involvement in RCTs (Figure 1). In addition, three systematic reviews included in the evidence synthesis addressed RCT integrity issues related to patient, public and consumer involvement.[27–29] This manuscript has been prepared in accordance with the GRIPP-2 guideline (Appendix 2).[30]

RESULTS

There were 30 stakeholders (Table 1) with voting rights from 14 countries in 5 continents including trialists, ethicists, methodologists, statisticians, consumer representative, industry representative, systematic reviewers, funding body panel members, regulatory experts, authors, journal editors, peer-reviewers and advisors for resolving integrity concerns. Their combined wide and appropriate expertise, based on

self-assessment, ranged broadly to include all aspects of the RCT research lifecycle from protocol development to knowledge transfer (Figure 1).

The initial long list of 111 statements (73 stakeholder-provided, 46 generated via evidence synthesis,[20] and 8 supported by both) was submitted to consensus via the modified Delphi survey (Figure 2). The first survey round had 26 out of 30 (86.7%) respondents and 64 statements were rated above the 76.5% APMO threshold for consensus. Among these, the strength of the agreement among stakeholders was good or very good in all the statements (Table 2). The remaining 47 statements along with the 7 new stakeholder-provided statements were subjected to revisions. After merging exact and inexact duplicates, 40 statements were submitted to the second survey round, where there were 26 out of 30 (86.7%) respondents and 24 statements were rated above the 68.4% APMO threshold for consensus. Among these, the strength of the agreement among stakeholders was good in 18 (75%) statements (Table 2). The 64 statements agreed in the first modified Delphi survey round were merged, removing exact and inexact duplications, to take forward 54 along with 24 agreed statements from second round to the consensus development meeting. The remaining 16 statements that lacked consensus after the second round were also taken forward. Sensitivity analysis for consensus threshold deploying the predefined arbitrary 70% cut-off showed that the APMO threshold was more conservative in the first round, permitting more statements to be re-examined (Table 2).

There was one new stakeholder-provided statement taking to total presented to 95 at this final stage. At the outset the stakeholder group confirmed that statements below 50% agreement threshold were to be excluded. Following discussion, merging, and voting in the consensus development meeting of the final shortlist contained 81 statements (49 stakeholder-provided, 41 systematic review-generated, 9 supported by both). Of the total, 32 (39.5%) were unique evidence-based statements. Of the 41 statements underpinned by evidence synthesis,[20] two were based on at least one high-moderate quality systematic review.[27,31] As shown in Table 3, the entire RCT lifecycle was covered with statements concerning general aspects (n=6), design and approval (n=11), conduct and monitoring (n=19), reporting of protocols and findings (n=20), post-publication concerns (n=12), and future research and development (n=13).

DISCUSSION

MAIN FINDINGS

Our international multi-stakeholder consensus provides the first specific integrity statement for promoting and protecting RCT integrity. It was developed in a robust and comprehensive manner, covering the entire RCT lifecycle. The general statements on RCT integrity emphasize the need for global harmonization and action. The statements relating to RCT design, approval, conduct and monitoring make clear that integrity needs embedding throughout the research lifecycle. The responsibilities of the publishing community are covered in statements concerning manuscript submission, peer-review, reporting and complaints. Further statements highlight the need for continuing research and development to advance responsible research conduct in RCTs. Drafted in a simple and clear language, the set of statements needs

implementation by the clinical trialist community and related institutions to take forward the health research integrity agenda.

LIMITATIONS AND STRENGTHS

There are several issues to consider in the weaknesses and strengths of this consensus development study. Defining research integrity to determine the statement scope was not straightforward. Although there is no agreed definition,[3,32] it is important to recognise that there is no controversy. To confidently use research results, society expects that the highest ethics standards and professionalism are deployed to conduct and report research.[1] Defining integrity narrowly, focusing on post-submission or post-publication dishonesty assessments, fails to recognize that the whole research journey needs addressing.[33] Our work is subject to other limitations including the possibility that the consensus group, which may be seen as having been derived from convenience sampling with a bit of snowballing, may not have included all perspectives despite an extensive effort to capture the widest possible range (Figure 1); our stakeholder group sample size was larger than the median of 22 experts included in previous reporting guideline development groups.[34] The surveys and voting were, by nature of the consensus, opinion-based. Not every stakeholder endorsed every statement (see percentages of agreement in Table 3). For example, despite the high level of overall support (92.3% approval with good level of agreement among stakeholders in the first round), there was a strong individual objection to the role of data monitoring committee in providing oversight for data integrity (Table 3, statement 26). In another example, where two statistics experts disagreed over the interpretation of the underlying evidence[35,36] used to formulate the statement

concerning statistical significance (Table 3, statement 33), the overall level of support just crossed the threshold for consensus (69.2% approval in the second round). For implementing this statement, examples of valid analytic strategies in the presence of multiple outcomes reported in the published literature can be helpful.[37–39] The use of the umbrella review[20] added breadth and objectivity.[40] For example, the statement concerning the input of professional medical writers arose from a systematic review (Table 3, statement 40).[20] It did not emerge from the input of any stakeholder. If a reader suspects a conflict of interest, we provide all the disclosures of stakeholders´ interests (Appendix 1). Another criticism may be that the stakeholders may have been too lenient, inclined to promote integrity softly, instead of creating challenges for researchers, committees, publishers, etc. through hard-to-implement recommendations. By explicitly reporting the agreement levels and openly sharing the consensus data we intended to maximize transparency for readers. The consensus statement would, no doubt, need updating and revisions in the future.

Our strength is that we captured integrity issues across the RCT lifecycle, advancing on previous general statements.[2,3] Using established, scientifically-based consensus techniques[21–25] we developed a specific statement that is comprehensive, methodologically replicable and transparently reported (see appendices concerning author contributions, disclosure statements, and data sharing). The umbrella review[20] contributed a high proportion of statements to those provided by stakeholders, who had a wide and appropriate range of expertise and experience including consumer representation.[41] It is important to note that stakeholders themselves were not authors of RCTs with active expressions of concerns or retractions related to integrity.

The lay member of the stakeholder group had experience of representing patients and public in research,[42] assisting trialists in design and conduct, serving as member of oversight committees, and scoring RCT grant applications for funding.

Surveys were anonymised with objectively determined statement approval thresholds and subjected to sensitivity analysis. Several statistics are available in the literature to determine the degree of consensus among respondents within a panel, including stipulated number of rounds, subjective analysis, APMO, mode, mean/median rating and others.[24] Our chosen statistics, APMO and the predefined arbitrary threshold, are among the most commonly used.[24] Additionally, we used IQR to quantify the strength of agreement among the stakeholders as a measure of how closely they agreed or disagreed with each other. The approval threshold was determined arbitrarily during the final voting round, something that should be improved in future consensuses. Through various consensus and feedback cycles, each statement was worded for maximum clarity of meaning and avoiding ambiguities. With focus on practicality, the statement set provides recommendation for embedding and enhancing RCT integrity standards. All the statements in the final set had high level of consensus across our stakeholder group.

INTERPRETATION OF THE FINDINGS

Our statement provides the basis for creating implementation plans and policies at stakeholder institutions and organisations to help inculcate integrity in RCTs. It is necessary to invest in the clinical research infrastructure required to support

trustworthy RCTs. Protecting and promoting RCT integrity requires a multifaceted approach, e.g. a combination of continuing education in best research practice in clinical trials targeting a range of audiences, improved governance and audit, and editor and peer-reviewer training in methodology. (Un)intentional errors can be reduced but cannot completely be eliminated. Admission of mistakes without the risk of persecution is a key aspect of continuous improvement.[43] To improve RCT credibility in health research, strategies to reduce the probability of errors are urgently required,[44] something that our statement emphasises. As far as trial oversight is concerned, the statement suggests that ethics committees, in addition to their traditional protocol appraisal and approval function before a trial can begin, should be given a role in monitoring the conduct of the trial. Deliberations of the trial oversight committees should be formally documented and, in the future, may need to be made publicly available during the course of the trial to match the growing transparency demands. On completion of the trial, chairs of ethics and oversight committees may provide certificates of authenticity to the authors for submission with their trials' manuscripts.

The statement recognises biomedical journals as key stakeholders in RCT integrity, as is obvious from the proportion of editors and peer-reviewers represented on our consensus group. It was recognised that majority of the journals' instructions to authors lacked sufficient detail to guide trialists to report their trial findings with integrity.[45] This was specifically highlighted to be the case for the information related to reporting of ethics approval, sources of finding, potential conflict of interests, trial registration and statistical analysis plans.[46–50] When an allegation of possible scientific

misconduct is made, journals have an obligation to investigate in an unbiased manner with an explicit policy about managing conflicts of interests of their editors, peer-reviewers and advisors. Our statement advises authors to actively engage with journal investigation process and submit their de-identifiable raw data to be examined if required. As a matter of good practice with respect to promoting transparency, authors can voluntarily electronically submit their data in a repository at the same time as submission of the trial manuscript. There is no logical reason to not be proactive, waiting for this to be made a mandatory requirement, which no doubt is the natural next step in the development of the ICMJE data sharing statement.[51] Hopefully, it will help in reducing the risk of complaints.

The reported prevalence of scientific misconduct is 2-14%.[52] During an investigation misconduct may appear obvious, for example when repeated duplications of observations (coping and pasting of rows and columns) or a formula to generate false data in a spreadsheet raise suspicion. However, in every case before arriving at a decision about flagging an RCT as being fraudulent a careful investigation of the raw data is required. If tools for detecting misconduct perform poorly, this would lead to false positive findings.[53] Wrongful accusations will damage science and healthcare.[43,54] Accurately detecting misconduct should therefore be a focus of future research to support peer-review and evaluation of post-publication concerns. Education in good research ethics, governance and monitoring may be currently more effective in generating trustworthy randomised evidence.[55,56]

CONCLUSION

Implementation of this international multi-stakeholder consensus will contribute to the enhancement of clinical trial integrity.

FUNDING

This research received no external funding. Travel expenses, accommodation and logistic facilitation of the consensus were by UEARS (Upper Egypt Assisted Reproduction Conference) 2022.

ETHICS/INSTITUTIONAL REVIEW BOARD STATEMENT

Not required

ACKNOWLEDGMENTS

K.S.K is a distinguished investigator at the University of Granada funded by the Beatriz Galindo (senior modality) program of the Spanish Ministry of Education; M.N-N is granted a research training fellowship by the Carlos III Research institute (Rio Hortega program - CM20/00074); Cairo consensus group would like to thank Upper Egypt Assisted Reproduction Conference (UEARS) for its support to this research integrity initiative; Cairo consensus group would like to thank COMSTECH, the Committee on Scientific and Technological Cooperation of a 57-country consortium, for its support to this research integrity initiative.

CONTRIBUTORS

Roles of authors are listed in Table 1 in accordance with the Contributor Role Taxonomy (CRediT).[57]

CONFLICTS OF INTEREST

The potential conflicts of interest for all the authors are listed in Appendix 1.

DATA SHARING STATEMENTS

A detailed description of results of each survey are openly available in Open Science Forum at https://osf.io/92ahr

REFERENCES

1. Steneck NH. Fostering integrity in research: definitions, current knowledge, and future directions. Sci Eng Ethics. 2006 Mar;12(1):53–74.
2. Resnik DB, Shamoo AE. The Singapore statement on research integrity. Account Res. 2011 Mar;18(2):71–5.
3. Moher D, Bouter L, Kleinert S, Glasziou P, Sham MH, Barbour V, et al. The Hong Kong principles for assessing researchers: Fostering research integrity. PLoS Biol. 2020 Jul 1;18(7).
4. Kretser A, Murphy D, Bertuzzi S, Abraham T, Allison DB, Boor KJ, et al. Scientific Integrity Principles and Best Practices: Recommendations from a Scientific Integrity Consortium. Sci Eng Ethics. 2019 Apr 15;25(2):327–55.
5. Di Renzo G C, Tosto V T V. The island of research (one rule): do not block the path of enquiry. In: Di Renzo G C Ed, editor. Essential Writing, Communication and Narrative Skills for Medical Scientists Before and After the COVID Era. Springer Nature, Basel; 2022. p. 1–17.
6. Departmental policy to guide integrity of science and scientific products.
7. European Network of Research Integrity Offices (ENRIO) [Internet]. [cited 2022 Jan 17]. Available from: http://www.enrio.eu/
8. World Conferences on Research Integrity Foundation (WCRIF) [Internet]. [cited 2022 Jan 17]. Available from: https://wcrif.org/
9. COPE: Committee on Publication Ethics | Promoting integrity in scholarly research and its publication [Internet]. [cited 2022 Jan 17]. Available from: https://publicationethics.org/

10. Bauchner H, Golub RM, Fontanarosa PB. Reporting and Interpretation of Randomized Clinical Trials. JAMA. 2019 Aug 27;322(8):732–5.

11. Hariton E, Locascio JJ. Randomised controlled trials – the gold standard for effectiveness research: Study design: randomised controlled trials. Vol. 125, BJOG: An International Journal of Obstetrics and Gynaecology. NIH Public Access; 2018. p. 1716.

12. de Vrieze J. Large survey finds questionable research practices are common. Science (80-). 2021;373(6552):265.

13. Gopalakrishna G, ter Riet G, Vink G, Stoop I, Wicherts JM, Bouter LM. Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in the Netherlands. PLoS One. 2022;17(2 February).

14. Steen RG, Casadevall A, Fang FC. Why Has the Number of Scientific Retractions Increased? PLoS One. 2013 Jul 8;8(7).

15. Resnik DB, Stewart CN. Misconduct versus honest error and scientific disagreement. Account Res. 2012 Jan;19(1):56–63.

16. Bolland MJ, Avenell A, Gamble GD, Grey A. Systematic review and statistical analysis of the integrity of 33 randomized controlled trials. Neurology. 2016;87(23):2391–402.

17. Guraya SY, Norman RI, Khoshhal KI, Guraya SS, Forgione A. Publish or perish mantra in the medical field: A systematic review of the reasons, consequences and remedies. Pakistan J Med Sci. 2016;32(6):1562–7.

18. Djurisic S, Rath A, Gaber S, Garattini S, Bertele V, Ngwabyt SN, et al. Barriers to the conduct of randomised clinical trials within all disease areas. Trials. 2017 Aug 1;18(1).

19. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH): Official web site [Internet]. [cited 2022 Jan 17]. Available from: https://www.ich.org/

20. Núñez-Núñez M, Maes-Carballo M, Mignini LE, Chien PF, Khalaf Y, Fawzy M, Zamora J, Khan KS B-CA. Research integrity in clinical trials: an umbrella review. BJOG. 2022;

21. Jandhyala R. Delphi, non-RAND modified Delphi, RAND/UCLA appropriateness method and a novel group awareness and consensus methodology for consensus measurement: a systematic literature review. Curr Med Res Opin. 2020;36(11):1873–87.

22. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs. 2000;32(4):1008–15.

23. Fitch K, Bernstein SJ, Mcdonnell J, Kahan JP. The RAND / UCLA Appropriateness Method User ' s Manual Appropriateness Method User ' s Manual Approved for Public Release Approved for Public Release. Transformation. 2001. 109 p.

24. Gracht HA Von Der. Technological Forecasting & Social Change Consensus measurement in Delphi studies Review and implications for future quality assurance. Technol Forecast Soc Chang. 2012;79(8):1525–36.

25. Cottam H-R, Roe M, Challacombe J, Roe M. Outsourcing of trucking activities by relief organisations. 2004;(January):1–26.

26. Shea BJ, Reeves BC, Wells G, Thuku1 M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised

studies of healthcare interventions, or both. BMJ. 2017;358:4008.

27. Houghton C, Dowling M, Meskell P, Hunter A, Gardner H, Conway A, et al. Factors that impact on recruitment to randomised trials in health care: a qualitative evidence synthesis. 2020;

28. Natale P, Saglimbene V, Ruospo M, Gonzalez AM, Strippoli GF, Scholes-Robertson N, et al. Transparency, trust and minimizing burden to increase recruitment and recruitment in trials: A systematic review. J Clin Epidemiol. 2021;134:35–51.

29. Paramasivan S, Davies P, Richards A, Wade J, Rooshenas L, Mills N, et al. What empirical research has been undertaken on the ethics of clinical research in India? A systematic scoping review and narrative synthesis. BMJ Glob Heal. 2021;6(5):1–19.

30. Staniszewska S, Brett J, Simera I, Seers K, Mockford C, Goodlad S, et al. GRIPP2 reporting checklists: Tools to improve reporting of patient and public involvement in research. Res Involv Engagem. 2017;3(1).

31. Marusic A, Wager E, Utrobicic A, Sambunjak D, Anderson MS, Rothstein HR. Interventions to prevent misconduct and promote integrity in research and publication. Cochrane Database Syst Rev. 2013;2013(2).

32. Kretser A, Murphy D, Bertuzzi S, Abraham T, Allison DB, Boor KJ, et al. Scientific Integrity Principles and Best Practices: Recommendations from a Scientific Integrity Consortium. Sci Eng Ethics. 2019 Apr;25(2):327–55.

33. Khan KS. Integrity culture is underpinned by education, not post-submission dishonesty assessments. Reprod Biomed Online. 2022;00(0):6483.

34. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. PLoS Med. 2010;7(2).

35. FDA guidance: Multiple Endpoints in Clinical Trials. Guidance for Industry. Jan 2017. FDA-2016-D-4460. Issued by the Center for Drug Evaluation and Research.

36. EMA guidance: POINTS TO CONSIDER ON MULTIPLICITY ISSUES IN CLINICAL TRIALS, Sep 2002, CPMP/EWP/908/99, Issued by the Committee for proprietary medicinal products (CMPM.

37. Schmid P, Adams S, Rugo HS, Schneeweiss A, Barrios CH, Iwata H, Diéras V, Hegg R, Im SA, Shaw Wright G, Henschel V, Molinero L, Chui SY, Funke R, Husain A, Winer EP, Loi S ELImTI. Atezolizumab and Nab-Paclitaxel in Advanced Triple-Negative Breast Cancer. N Engl J Med. 2018;379(22):2108–21.

38. Smith MR, Hussain M, Saad F, Fizazi K, Sternberg CN, Crawford ED, Kopyltsov E, Park CH, Alekseev B, Montesa-Pino Á, Ye D, Parnis F, Cruz F, Tammela TLJ, Suzuki H, Utriainen T, Fu C, Uemura M, Méndez-Vidal MJ, Maughan BL, Joensuu H, Thiele S, Li R, Kuss I TBATI. Darolutamide and Survival in Metastatic, Hormone-Sensitive Prostate Cancer. N Engl J Med. 2022;386(12):1132–42.

39. Fizazi K, Foulon S, Carles J, Roubaud G, McDermott R, Fléchon A, Tombal B, Supiot S, Berthold D, Ronchin P, Kacso G, Gravis G, Calabro F, Berdah JF, Hasbini A, Silva M, Thiery-Vuillemin A, Latorzeff I, Mourey L, Laguerre B, Abadie-Lacourtoisie S, Martin E BAP-1 investigators. Abiraterone plus prednisone added to androgen deprivation therapy and docetaxel in de novo metastatic castration-sensitive prostate cancer (PEACE-1): a multicentre, open-label, randomised, phase 3 study with a 2 × 2 factorial design. Lancet (London, England). 2022;399(10336):1695–707.

40. Qaseem A, Forland F, Macbeth F, Ollenschläger G, Phillips S van der WP. Guidelines International Network: toward international standards for clinical practice guidelines. Ann Intern Med. 2012;156(7):525–31.

41. García-Martín M, Amezcua-Prieto C, H Al Wattar B, Jørgensen JS, Bueno-Cavanillas A, Khan KS. Patient and public involvement in sexual and reproductive health: Time to properly integrate citizen's input into science. Vol. 17, International Journal of Environmental Research and Public Health. 2020. p. 1–12.

42. Moss N, Daru J, Lanz D, Thangaratinam S, Khan KS. Involving pregnant women, mothers and members of the public to improve the quality of women's health research. Vol. 124, BJOG: An International Journal of Obstetrics and Gynaecology. 2017. p. 362–5.

43. Macleod M. Want research integrity? Stop the blame game. Nature. 2021;599(7886):533.

44. Nolan TW. System changes to improve patient safety. Br Med J. 2000;320(7237):771–3.

45. Malički M, Jerončić A, Aalbersberg IjJ, Bouter L, ter Riet G. Systematic review and meta-analyses of studies analysing instructions to authors from 1987 to 2017. Vol. 12, Nature Communications. 2021.

46. Malički M, Jerončić A, Aalbersberg IjJ, Bouter L, ter Riet G. Systematic review and meta-analyses of studies analysing instructions to authors from 1987 to 2017. Vol. 12, Nature Communications. 2021.

47. Goldstein CE, Weijer C, Brehaut JC, Fergusson DA, Grimshaw JM, Horn AR, et al. Ethical issues in pragmatic randomized controlled trials: A review of the recent literature identifies gaps in ethical argumentation. BMC Med Ethics. 2018;19(1):1–10.

48. Schellings R, Kessels AG, ter Riet G, Knottnerus JA, Sturmans F. Randomized consent designs in randomized controlled trials: Systematic literature search. Contemp Clin Trials. 2006;27(4):320–32.

49. Darmon M, Helms J, De Jong A, Hjortrup PB, Weiss E, Granholm A, et al. Time trends in the reporting of conflicts of interest, funding and affiliation with industry in intensive care research: a systematic review. Intensive Care Med. 2018;44(10):1669–78.

50. Bekelman JE, Gross CP. Scope and impact of financial conflicts of interest in biomedical research: A systematic review. JAMA - J Am Med Assoc. 2003;289(4):454–65.

51. International Committee of Medical Journal Editors' (ICMJE) [Internet]. Available from: http://www.icmje.org/disclosure-of-interest/

52. Fanelli D. Why Growing Retractions Are (Mostly) a Good Sign. PLOS Med. 2013;10(12):e1001563.

53. Hartgerink C, Voelkel J WJ. Detection of data fabrication using statistical tools. PsyArXiv. 2019;1–50.

54. Lewandowsky S, Bishop D. Research integrity: Don't let transparency damage science. Nature. 2016;529(7587):459–61.

55. Khan KS. 'Flawed use of post publication data fabrication tests''. Research misconduct tests: putting patients' interests first".' J Clin Epidemiol. 2021;138(xxxx):227.

56. Khan KS. Assessing Research Misconduct in Randomized. Obstet Gynecol. 2021;138(6):944.

57. Allen L, O'Connell A K, V. How can we ensure visibility and diversity in research contributions? How the Contributor Role Taxonomy (CRediT) is helping the shift from authorship to contributorship. Learn Publ. 2019;32(1):71–4.

Figure 1: Expertise and experience of the stakeholder group in the international multi-stakeholder consensus statement on clinical trial integrity.

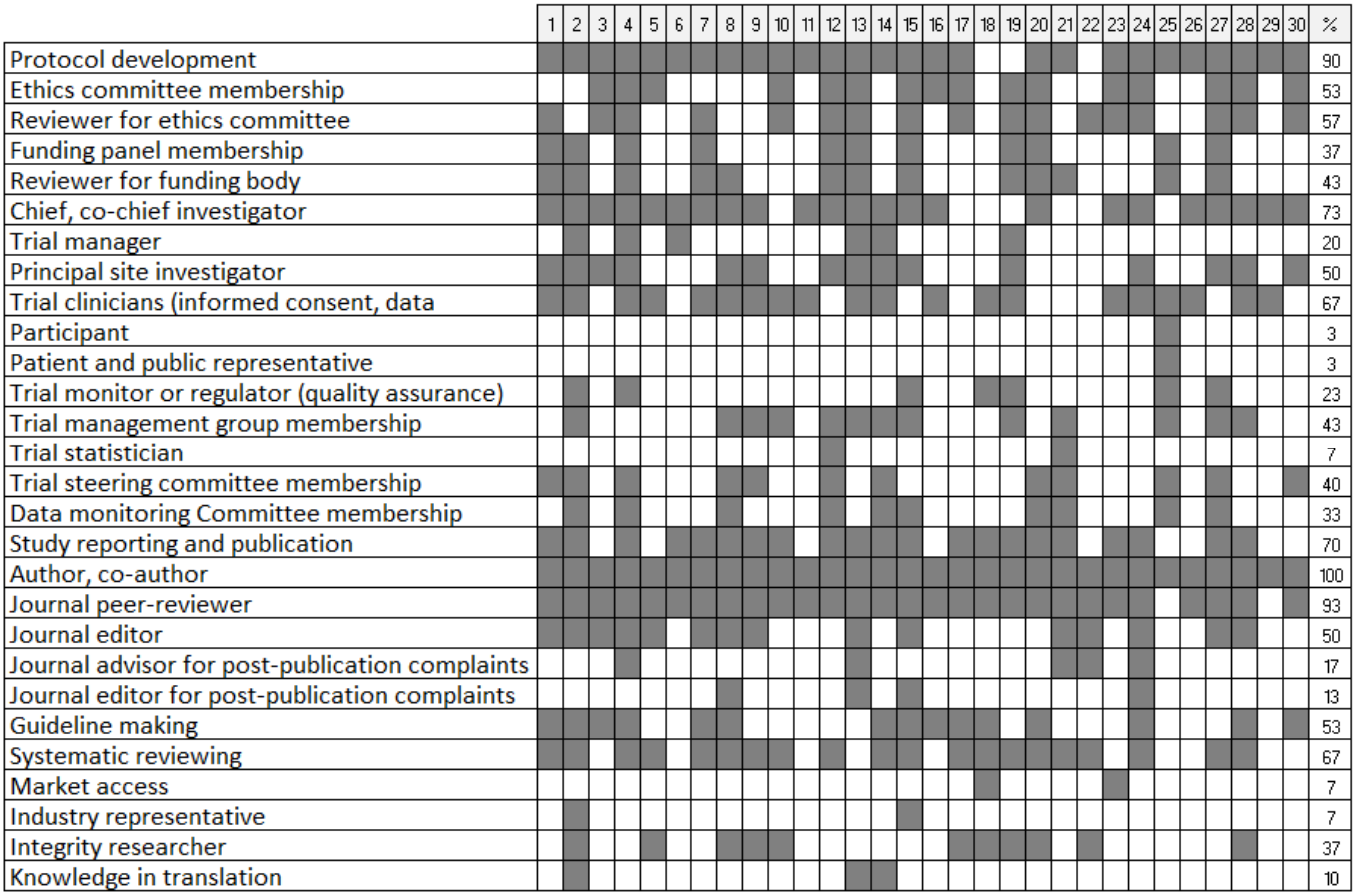

Figure 2: Flowchart of the development process for the international multi-stakeholder consensus statement on clinical trial integrity.

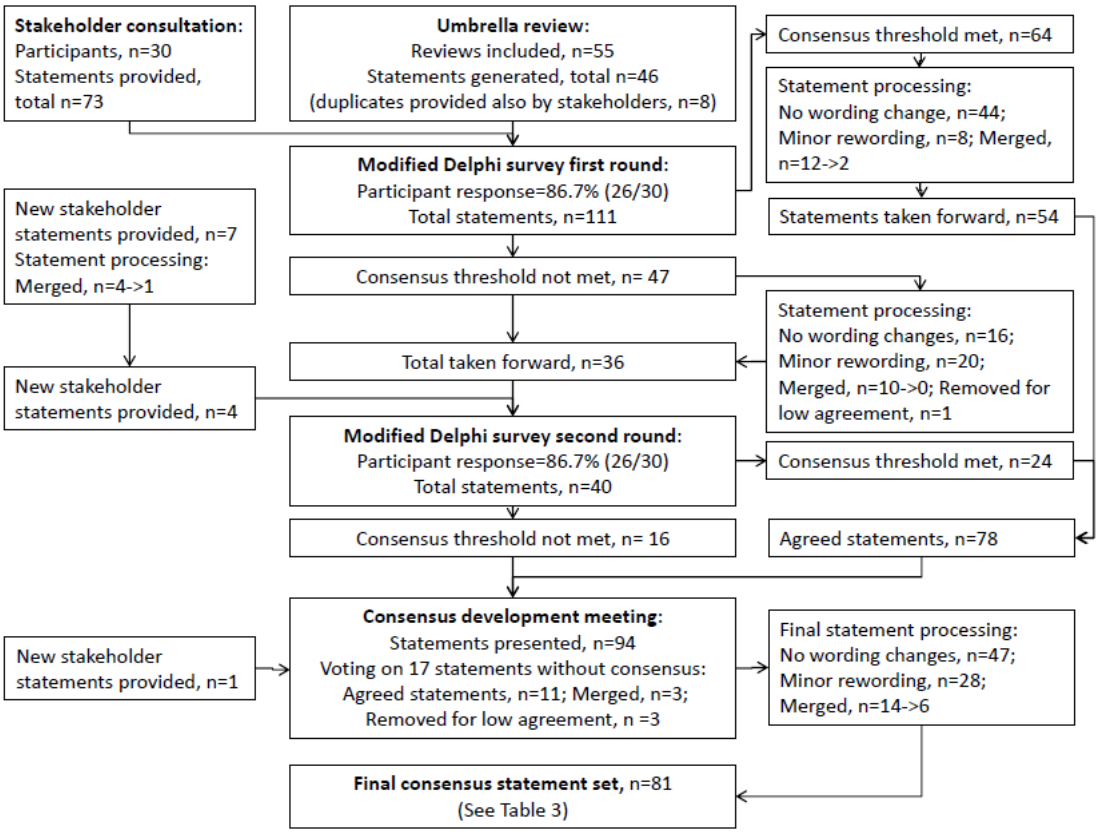

Table 1. Roles and filiation of the stakeholder group in the international multi-stakeholder consensus statement on clinical trial integrity.

Name	Role(s) of the authors	Affiliation	ORCID ID
Yacoub Khalaf	Conceptualization, convener, supervision, scientific contribution, review and editing and stakeholder	Guy's & St Thomas' Hospital Foundation Trust, UK	0000-0002-5642-7367
Khalid Saeed Khan	Conceptualization, convener, supervision, scientific contribution, review and editing and stakeholder	University of Granada; CIBERESP. Spain	0000-0001-5084-7312
Mohamed Fawzy	Conceptualization, methodology, project administration, scientific contribution and stakeholder	IbnSina, Banon Amshaj and Qena IVF Centres, Egypt	0000-0001-8756-3612
Patrick Chien	Scientific contribution, validation, writer, review and editing, and stakeholder	RUMC, Penang, Malaysia	0000-0002-5998-9592
Aurora Bueno-Cavanillas	Scientific contribution, methodology, validation and stakeholder	University of Granada; CIBERESP. Spain	0000-0002-0649-3016
Maria Nunez-Nunez	Writer, data curation and stakeholder	San Cecilio University Hospital; Ibs Granadada; CIBERESP. Spain	0000-0002-2633-4207
Marta Maes-Carballo	Writer, data curation and stakeholder	Complexo Hospitalario de Ourense; Hospital Público Verín. Spain.	0000-0002-4852-5100
Gamal Serour	Scientific contribution, validation, representative of EFSS and stakeholder	Al-Azhar University and Egyptian IVF-ET Centre, Egypt	0000-0002-0067-7850
Mohamed Aboulghar	Scientific contribution, validation, representative of MEFS and stakeholder	Cairo University and Egyptian IVF-ET Centre, Egypt	0000-0002-3935-6501
Gerben ter Riet	Scientific contribution, validation and stakeholder	Amsterdam University, Netherlands	0000-0002-2231-7637
Javier Zamora	Scientific contribution, statistician, writer and stakeholder	Hospital Ramón y Cajal, IRYCIS. Madrid, Spain and Birmingham University, UK	0000-0003-4901-588X
Jeffery Andrews	Scientific contribution and stakeholder	BD Integrated Diagnostic Solutions, USA	0000-0003-2416-0490
Hassan Sallam	Scientific contribution, and representative of ERC-RCOG and stakeholder	Alexandria University, Egypt	0000-0003-1308-6280
Jack Wilkinson	Scientific contribution and stakeholder	Centre of Biostatistics, Manchester, UK	0000-0003-3513-4677
Hazem Abdelghaffar	Scientific contribution and stakeholder	Sohag University, Egypt	Not available
Jacek Walczak	Scientific contribution and stakeholder	Centre of Excellence in Systematic Reviews, Central and Eastern Europe, CERTARA, Poland	0000-0003-4965-0461
Tayyiba Wasim	Scientific contribution and stakeholder	Services Institute of Medical Sciences, Services Hospital, Lahore, Pakistan	0000-0003-2444-9817
Ngawai Moss	Scientific contribution and stakeholder	University of London, UK	0000-0001-9369-5072
Hassan Maghraby	Scientific contribution, EFRE representative and	Alexandria University, Egypt	0000-0003-3661-1594

	stakeholder		
Jun Jim Zhang	Scientific contribution and stakeholder	Shangai Jiao Tong University School of Medicine, Shangai, China	Not available
Ali Mahran	Scientific contribution and stakeholder	Assiut University, Egypt	0000-0001-7870-4110
Luciano Mignini	Scientific contribution and stakeholder	Hospital Escuela Eva Perón de Granadero Baigorria; Grupo Oroño. Argentina	0000-0002-7783-9088
Mahmoud Abdelaleem	Scientific contribution and stakeholder	Assiut University, Egypt	0000-0003-3942-9325
Mohamed Bedaiwy	Scientific contribution and stakeholder	University of British Columbia, Canada	0000-0002-3454-8555
Chris Hartgerink	Scientific contribution and stakeholder	Liberate Science GmbH, Germany	0000-0003-1050-6809
Mohamed Sabry	Scientific contribution and stakeholder	Sohag University, Egypt	0000-0002-8206-2074
Mohamed Yahya AbdelRahman	Scientific contribution and stakeholder	Sohag University, Egypt	0000-0002-0136-512x
Gian Carlo Di Renzo	Scientific contribution and stakeholder	University of Perugia, Perugia, Italy	0000-0003-4467-240X
Zahida Qureshi	Scientific contribution and stakeholder	University of Nairobi, Kenia	0000-0003-4223-3227
Abdullah Alkhenizan Alkhenizan	Scientific contribution and stakeholder	Al Faisal University, Saudi Arabia	0000-0002-0269-5200
David Mortimer	Advisor, consensus methodology and statement wording	University of Dundee, Scotland, UK and Oozoa Biomedical Inc, Canada	0000-0002-0638-2893

Table 2. Statements reaching consensus according to the different approval thresholds for agreement in the multi-stakeholder international consensus concerning clinical trial integrity.

Analysis	Number of agreed statements (%)	
	1st round survey (Total=111)	2nd round survey (Total=40)
Main analysis[a]		
Above APMO approval threshold	64 (57.7%)	24 (60.0%)
Strength of agreement among stakeholders concerning statements above APMO threshold [b]		
IQR 0 (very good)	4/64 (6.2%)	0/24 (0%)
IQR 1 (good)	60/64 (93.8%)	18/24 (75.0%)
IQR ≥2 (poor)	0/64 (0%)	6/24 (25.0%)
Sensitivity analysis[c]		
Above predefined arbitrary approval threshold	74 (67.6%)	17 (42.5%)

a. APMO: Average Percent of Majority Opinions. In this computation, a statement was considered as agreed if the majority (>50%) of stakeholders responded "strongly agree" or "agree" on the seven-point scale. A statement was considered as disagreed if the majority (>50%) of stakeholders responded "disagree" or "strongly disagree" on the seven-point scale. The AMPO approval threshold was calculated as: sum of majority agreed and majority disagreed statements / total number of responses received x 100%. APMO approval thresholds were 76.4% in Delphi 1st round and 68.4% in Delphi 2nd round.

b. IQR: Interquartile range of the responses in the seven-point scale. In this computation, IQR 0 (>50% stakeholders gave the same responses) indicated very good strength of agreement; IQR 1 (>50% stakeholders range of responses was ≤2 points of the scale) indicated good strength of agreement; IQR ≥2 (>50% stakeholders gave responses >2 points of the scale) indicated poor strength of agreement.

c. Predefined arbitrary approval threshold was >70%.

Table 3. Statements concerning clinical trial integrity from a multi-stakeholder international consensus (n=81).

Final consensus statements	Agreement (%)*			Underpinning information source**
	Delphi 1st round (threshold 76.5%)	Delphi 2nd round (threshold 68.4%)	Consensus meeting	
General				
1. Clinical trial integrity guidelines and policies must be explicit, visible, and prospectively enforceable at all levels through an implementation plan.	82.7[a]			SPS
2. Trialists, ethics committee members, journals editors and peer-reviewers should receive appropriate methodological and integrity training.	80.8[a]			SPS,1-7
3. Trial ethics committees should have accreditation and regional, national and international harmonisation of ethics assessment criteria and review process.	92.3[a]			8,9
4. There should be continuous public documentation of trials during the entire study lifecycle.	61.5	61.5	80.0	SPS
5. Journals should support adoption of responsible research practices in the design, conduct, analysis, reporting and archiving of trials.	88.5			SPS
6. Institutions should avoid excessive publication pressure.	76.9			SPS

Design and approval				
7. Ethics approval should be obtained for all trials, including those using de-identified data.	67.3[a]	65.5[a]	100	10,11,20,21
8. Informed consent should be developed with patient (or their representative) and public involvement.	80.8			12,13,14,15,16
9. Informed consent should be examined and approved by the ethics committee.	96.2			1,12,14
10. Informed consent should include explicitly how the de-identified data will be shared at the time of publication or used for future analysis	73.1	65.4	96.4	17
11. Trials should be prioritised and resourced according to local health care needs, strategy, and culture, especially in multi-country trials including low-resource settings.	69.2	69.2[f]		1,12,18
12. Trials should be approved according to local ethics and regulatory framework, especially in multi-country trials including low-resource settings.	76.9			1,12,18
13. Translations of patient reported outcomes should be culturally sensitive in multi-country trials including low-resource settings.	84.6			19
14. Equality, diversity and inclusion should be embedded in trial design to maximize generalisability of findings.	76.9			SPS
15. Sample size estimation should be sufficiently detailed to permit replication.	92.3			24
16. Primary and secondary outcomes should follow the internationally agreed core outcomes	80.8			SPS

whenever available.				
17. The trial protocol, including ethics approval, should be prospectively registered with an open-access trial registry prior to participant recruitment. This policy should be included in research institutions` and sponsors` regulations, and researcher employment and funding contracts.	78.9[a,b,f]			SPS, 30,32,35
Conduct and monitoring				
18. Trial site assessment should put in place measures to mitigate integrity breaches with the support of local research governance departments.	88.5[a]			SPS
19. There should be promotion of admission of honest or unintentional errors in the conduct of the trial without fear of blame. A part of this policy should be training.	94.2[a]			SPS
20. Innovative recruitment strategies should be participant-driven and should comply with ethics principles.	88.5			15,25,26[e]
21. Routinely collected data should be validated before analysis and reporting.	69.2	84.6		SPS, 20,27
22. Informed consent oversight should be part of trial audit.	92.3			10,13
23. The membership of independent trial steering and data monitoring committees should declare any potential conflict of interests.	100			SPS
24. The membership of independent trial steering committees should include patient and public stakeholders.	69.2	65.4	79.3	SPS
25. Minutes of the independent trial steering and data monitoring committees should be available	69.2	61.5	83.0	SPS

when required.				
26. Data monitoring committee charter should include responsibility for data integrity.	92.3			SPS,28
27. Centralized monitoring and selective source data verification should be deployed for ensuring data integrity.	80.8			29
28. There should be transparency in the method(s) of handling missing data at all stages of monitoring and reporting.	96.2			SPS
29. Early termination of a trial should be undertaken with the input of the independent trial steering and data monitoring committees.	96.0			SPS
30. Any amendment to study protocol should be reported to the trial registry (with dates). Major changes also require ethics approval.	100			SPS
31. The statistical analysis plan should be developed and published at the start or during the early stages of the trial before the data is made available to the investigators.	88.5			SPS
32. All analyses should be pre-specified from the outset (the analysis of the primary outcome and secondary outcomes, sub-group analyses, and sensitivity analyses).	84.6			SPS
33. There should be a single primary outcome pre-specified; when there are multiple key outcomes, valid testing strategies should be considered for maintaining familywise type-1 error within the acceptable limit of 5 %.	65.4	69.2[f]		SPS
34. Trial funders should mandate in their contract with researchers that outcomes are analysed and reported according to preregistration.	42.3	57.7	88.0	SPS

35. Databases for trials should include auditable access logs and permission management systems to prevent illicit access to data or editing of data.	n/a[d]	n/a[d]	100	SPS
36. Trial integrity and quality evidence synthesis both require the avoidance or minimisation of bias in trial conduct.	n/a[d]	84.6		SPS
Reporting of protocols and findings				
37. Trialists are strongly encouraged not to submit to a predatory journal, avoiding journals without transparency and integrity.	69.2	65.4[a]	83.3	30
38. Journals' authors' instructions should explicitly and comprehensively cover the requirements for openness and transparency.	84.6[a]			SPS, 31,32,33,34
39. Journals´ electronic submission system should facilitate compliance with the integrity-related authors` instructions.	73.1	92.3		SPS
40. Professional medical writing could help in reporting more clearly and succinctly to meet the integrity requirements. Its contribution should be reported.	61.5	69.2[f]		36
41. The speed with which editorial and peer-review decisions are made should be balanced against the possibility of future complaints and retraction.	65.4	65.4	83.3	37
42. Reporting of ethics approval and informed consent details should be obligatory part of reporting guidelines and authors' instructions.	84.6[a]			10,13, 14,17,38
43. Ethics or independent data monitoring committee should provide confirmation that the trial was	61.6[a]	69.5[a]		SPS

conducted as planned.				
44. Authorship contribution (credit according to international guidelines) should be made explicit in the manuscript.	94.3[a]			SPS,22,23
45. Trial protocol and statistical analysis plan should be submitted in unredacted form along with data set, statistical syntax and analytical outputs.	69.2	88.5		SPS,7,33
46. Reporting of conflict of interests, funding sources and payments received by all authors should be standardised.	78.9[a]			SPS,23,34,39,40,41
47. Declaration of conflict of interest, funding sources and payments should be mandatory for peer-reviewers and editors.	88.5			SPS
48. Reporting of patient and public involvement in the trial should be mandatory.	76.9			SPS
49. Manuscripts should be prepared according to standard reporting guidelines (e.g SPIRIT, CONSORT, GRIPP-2, etc) and their specific extensions for particular trial types (e.g. human challenge trials, trials of social and psychological interventions, etc.).	76.9[a,c,f]			SPS,42,43, 47
50. Plagiarism checks should be routinely carried out on the article main text.	84.6			44
51. Errors, deviations from protocol, losses to follow-up, missing outcome data and solutions applied should be transparently reported.	92.3			45,46,54
52. Reporting the use of data monitoring committees, its responsibilities and its membership should be mandatory.	73.1	96.2		28

53. Among trials conducted in various languages use of translations in patient reported outcomes should be explicit.	53.8	53.8	91.6	19
54. Primary and secondary outcomes should be mandatorily linked to prospectively registered outcomes.	76.9			35
55. Spin in writing to misrepresent, overinflate or distort the methods, findings, results and conclusions should be eliminated.	82.7[a]			SPS
56. The strengths and limitations of the integrity-related issues, as well as any flaws in terms of less-than-ideal method implementation that was unavoidable, should be discussed in the manuscript.	73.1	96.2		SPS
Post-publication				
57. When a post-publication review detects integrity breaches, the implication is that the scientific process failed, so the focus should be on correction and learning lessons openly and collectively.	76.9			SPS
58. Journals have the responsibility to conduct their pre-publication assessments and peer-review in a manner so as to minimise the risk of post-publication dishonesty allegations.	92.3			SPS
59. Any guidance concerning post-publication integrity concerns (e.g, COPE https://publicationethics.org, https://doi.org/10.24318/o1VgCAih, https://doi.org/10.24318/cope.2019.2.4) should explicitly emphasise the investigators` responsibility to evaluate the integrity of the complaint and to support the trialists.	73.1	88.5		SPS
60. Institutions and journals should be equally supportive to the complainant(s) and author(s) in handling such complaints. There is a responsibility to protect honest trialists against harassment.	84.6[a]			SPS

61. Trialists must engage with any request for an explanation for apparent data discrepancy if required by the journal during both peer-review and post publication stages, or by systematic reviewers during evidence synthesis.	92.3			SPS
62. Trialists have the responsibility to keep detailed records of their trial including original protocol (with any subsequent amendments), ethics approval, details of the trial registration, de-identified raw data set, randomisation sequence employed, statistical plan, syntax and outputs of all the statistical analyses in case these are required to address any post-publication complaints.	80.8			SPS
63. Declaration of conflicts of interest, funding sources and payments should be mandatory for complainants.	84.6			SPS
64. Journals should act in an unbiased fashion transparently managing the conflict of interest of their own editors and advisors handling complaints.	80.8[a]			SPS
65. Trialists, with their institutional input, should be permitted to provide independent expert reports to the journal investigating a complaint.	76.9			SPS
66. If honest mistakes are identified in post publication, an erratum should be published.	96.2			SPS
67. Retraction notices should be clear and interpretable.	88.5			48
68. Post-retraction management of trials with proven misconduct should be based on a system that avoids continued citation and data misuse.	96.2			48
Future research and development				

69. Educational effectiveness of integrity training should be evaluated.	69.2	84.6		53[e]
70. The factors influencing participant willingness to give consent for data sharing should be evaluated.	61.5	76.9		51,52
71. The minimum requirement for adequate informed consent should be established.	61.5	69.2		49
72. The criteria for and level of data auditing required during conduct of trial should be delineated.	61.5	65.4	100	10,49
73. The integrity remit of data monitoring committees should be clarified.	69.2	80.8		28
74. The best method(s) for publication credit (authorship contribution) should be determined.	65.4	88.5		50
75. Effective peer review models should be developed for evaluation of trials.	84.6			55
76. Automated checks for compliance with reporting guidelines items (e.g CONSORT, SPIRIT, GRIPP-2) should be developed.	80.8			SPS
77. For the raw data to be shared, journals should clarify the requirements, e.g. randomisation sequence, cleaned or original de-identified dataset, statistical codes, etc.	69.3[a]	92.3		SPS
78. The validity of early post-submission and post-publication integrity tests should be evaluated.	65.4	84.6		44
79. A common research terminology should be developed for prevention of selective reporting.	57.7	53.8	86.9	54
80. Evidence syntheses of trials using reported study-level (not raw) data should develop methods	n/a[d]	69.2[f]		SPS

(e.g. subgroup meta-analyses or meta-regression) to evaluate integrity concerns.				
81. Evidence syntheses of trials should develop methods to access patient-level (raw) data to maximize transparency.	n/a[d]	76.9		SPS

For more details see Figure 2 and data sharing file (https://osf.io/92ahr)

* Agreement (%) for the Delphi rounds is the percentage of the sum of the "strongly agree" and "agree" responses provided on the seven-point scale for the approval of each individual statement by the stakeholders. Agreement (%) for the consensus meeting is the percentage of votes casted in favour of the total votes.

**List of references is provided in Appendix 3; SPS: Statement provided by stakeholders.

a. Median agreement (%) is shown for several merged statements.

b. The agreement percentage (78.9 %, the median of 88.5%, 84.6%, 73.08% and 61.54%) represents data for a merged statement containing four statements, two approved in the first round (related to prospective registration, 88.5% and 84.6%) and the other two approved in the second round (related to the policy, 73.08 % and 61.54% in the first round and they passed the approval threshold in the second round with 80.77% and 69.23%). The strength of agreement among stakeholders for those statements approved in second round was poor in the first round and good/poor in the second round (see methods and Table 2 for details).

c. The agreement percentage (76.9%, the median of 84.6% and 69.2%) represents data for a merged statement containing two statements, one approved in the first round (related to standard reporting guidelines, 84.6%) and the other approved in the second round (related to specific extensions, 69.2% in the first round and it passed the approval threshold in the second round with 69.2%). The strength of agreement among stakeholders for the specific extensions statement was good in the first round and poor in the second round (see methods and Table 2 for details).

d. n/a means not applicable, statement was provided by a stakeholder after the first or the second Delphi rounds.

e. Systematic review classified as "high" to "moderate" quality according to modified AMSTAR-2 ("Research integrity in clinical trials: an umbrella review. Reference in press")

f. Strength of agreement among stakeholders poor (see methods and Table 2 for details)