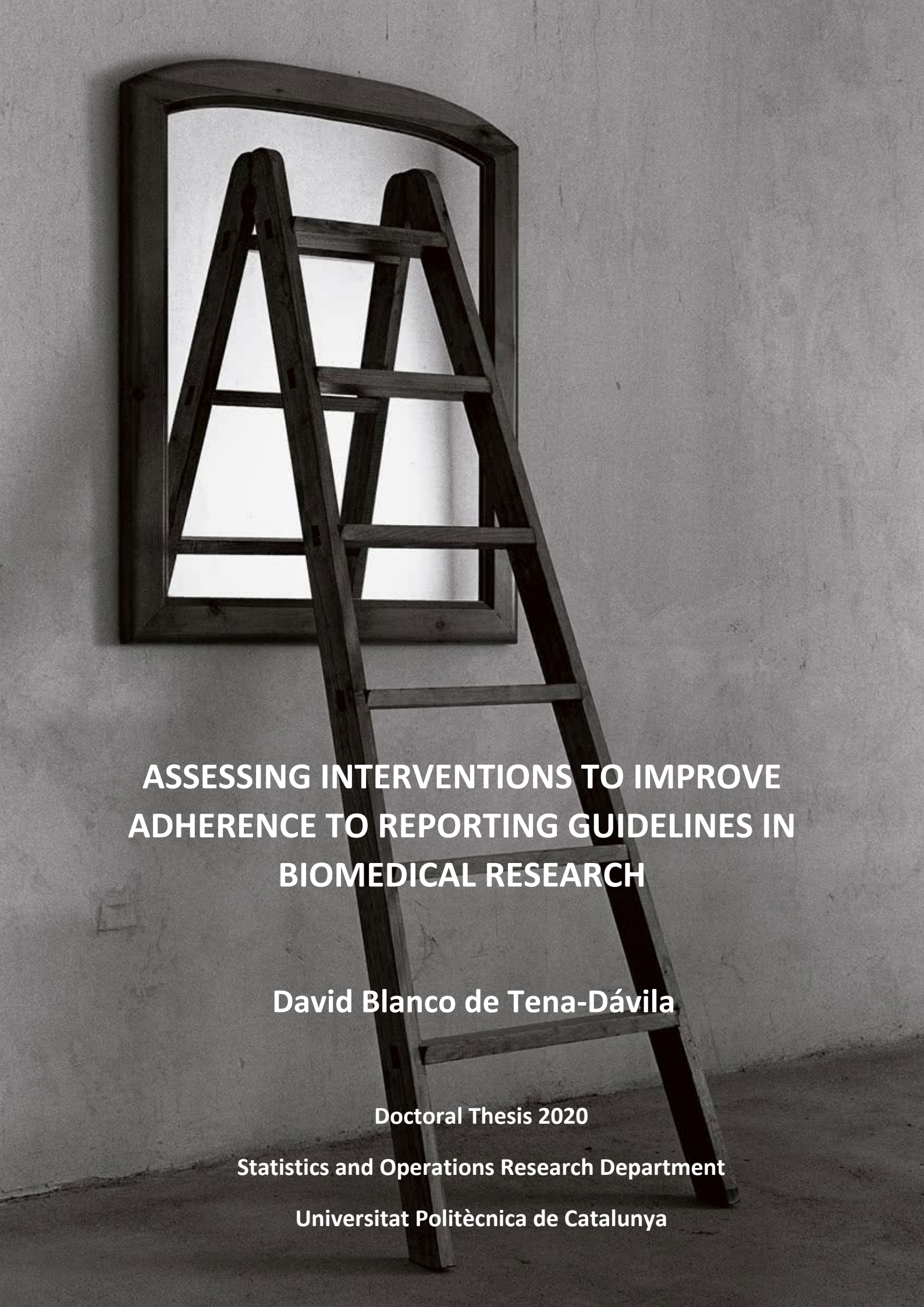


ASSESSING INTERVENTIONS TO IMPROVE ADHERENCE TO REPORTING GUIDELINES IN BIOMEDICAL RESEARCH

David Blanco de Tena-Dávila

Doctoral Thesis 2020

Statistics and Operations Research Department

Universitat Politècnica de Catalunya

Assessing interventions to improve adherence to reporting guidelines in biomedical research

PhD in Biostatistics and Operations Research

Statistics and Operations Research Department, Universitat Politècnica de Catalunya

PhD in Biostatistics

CRESS, INSERM, and INRA, Université de Paris

Doctoral thesis by

David Blanco de Tena-Dávila

Thesis supervisor

Erik Cobo (*Universitat Politècnica de Catalunya)*

Thesis co-supervisors

Jamie J Kirkham (*University of Manchester*)

Isabelle Boutron (*Université de Paris*)

Barcelona, May 2020

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 676207.

Agradecimientos (español)

Nunca hubiera imaginado que el apartado de agradecimientos se convirtiese en la parte más difícil de escribir de mi tesis. En estos tres años y medio que han transcurrido desde que llegué a una Barcelona que solo existía en mi imaginación, nutrida por las novelas de Eduardo Mendoza, he conocido a muchas personas sin las cuales ahora difícilmente estaría escribiendo estas líneas. La necesidad de mostrarles mi gratitud a todas estas personas, y a algunas otras, se hace más urgente después de cuarenta y siete días confinado en soledad. Espero que muy pronto podamos encontrarnos y que pueda encontrar un modo más efectivo de expresarles mi gratitud.

En primer lugar, quiero darle las gracias a Erik, mi director, por apostar por mí pese a mi falta de experiencia inicial y por guiarme durante este largo viaje. De su mano me incorporé a MiRoR. Me siento profundamente orgulloso de haber participado y contribuido al desarrollo de este fascinante proyecto. Actividades como los eventos de formación y de discusión de artículos o el proyecto común me han permitido explorar territorios desconocidos y acceder a una formación integral y crítica en el campo de la meta-investigación. Me gustaría expresar también mi gratitud a Isabelle Boutron por liderar esta iniciativa y, más personalmente, por ayudarme a tomar decisiones importantes cada vez que me encontraba en una encrucijada. Estoy especialmente agradecido a Jamie J. Kirkham, mi co-supervisor, por haber participado activamente en los distintos proyectos que integran esta tesis y por estar siempre presente a la hora de debatir el progreso de esta tesis. También quiero dar las gracias a otros colaboradores e integrantes de MiRoR, como David Moher o Douglas Altman, que me enseñaron a disfrutar de la investigación y a otorgarle más importancia a la calidad, la relevancia y el

rigor en la actividad científica que a la mera productividad. La desgraciada muerte de Doug hace poco menos de dos años me hará conservar con especial cariño los momentos compartidos con él.

Gracias también al grupo editorial BMJ Publishing Group y a Adrian Aldcroft, editor jefe de *BMJ Open*, por ofrecernos la oportunidad de realizar un ensayo aleatorizado en un entorno editorial real. Y gracias, muy especialmente, a Sara Shroter; sin su aportación y, sobre todo, sin su cercanía, profesionalidad y amabilidad, esta historia hubiera sido radicalmente distinta.

Gracias a mis compañeros de doctorado en la UPC por hacer el día a día más fácil. Nuestros debates y reuniones académicas nos han hecho más críticos con nuestro trabajo. Además, el apoyo mutuo nos ha fortalecido en momentos donde derrumbarse parecía la única opción. Mención especial a las tardes de amistad, cultura y bravas con mi compañera Yovi, cruciales para no volverme loco durante el intenso último año de doctorado. También quiero agradecer la labor, tan invisible como imprescindible, de Sonia, que ha dado la cara por nosotros cada vez que algún problema burocrático aparecía. Por último, gracias a los compañeros del Departamento de Estadística e Investigación Operativa y del GRBIO, muy especialmente a Lupe, Klaus y José Antonio por echar un cable siempre que lo necesité.

Además, quiero dar las gracias a mis amigos de siempre de Madrid, y a los que fui dejando atrás en Stuttgart y Liverpool. Por suerte, la telecomunicación nos acerca, y los muchos reencuentros me permiten albergar la esperanza de que ni el tiempo ni la distancia podrán con todo lo que hemos construido. Gràcies també a tu, Josep, per haver

entés les meves circunstàncies des que vaig arribar a Barcelona i per donar-me suport i un cop de mà sempre que ha calgut.

En último lugar, gracias a mi familia. En especial, gracias a mis padres, mi hermana y mis abuelos, que han confiado en mí y me han dado cariño y apoyo no solo durante estos años, sino durante toda mi vida. Sin ellos, esta historia nunca hubiera comenzado.

Acknowledgements (English)

I would have never imagined that writing the acknowledgements section would become the most difficult part of my thesis. It has been three years and a half since I arrived in Barcelona, a city that I had built in my imagination based on Eduardo Mendoza's novels. During this time, I have met many people that have made this thesis possible. The need to show my gratitude to all these people is even more urgent after I have spent 47 days confined alone because of a global pandemic. I hope we can meet soon.

First of all, I would like to thank Erik, my supervisor, for his support and guidance during this journey. Thanks to him I joined MiRoR. I feel deeply honoured to have been part of this project and to have contributed to its success. Activities such as the training events, the journal clubs, and the common project gave me the opportunity to explore in depth different topics within the field of meta-research. Thanks to Isabelle Boutron for leading this initiative and for helping out every time I was at a crossroad. I am also very grateful to my co-supervisor Jamie J Kirkham, who always found the time to provide feedback and discuss every step of my thesis. I also appreciate the collaboration of other MiRoR members such as David Moher or Douglas Atlman, who taught me the joy of doing research and the importance of focusing on the quality, relevance and rigour, rather than the quantity, of our scientific output. The sad loss of Doug almost two years ago makes me appreciate even more the moments we shared with him.

Thanks to BMJ Publishing Group and to Adrian Aldcroft, the editor-in-chief of BMJ Open, for giving us the chance to perform a randomised trial in a real editorial context. Special thanks to Sara Schroter; without her input and, above all, her closeness, dedication and kindness, this experience would have been radically different.

Thanks to my PhD colleagues at UPC for making everyday life easier. Our debates and academic meetings have made us more critical of our work. In addition, we have constantly supported each other, which has strengthened us at times when collapsing seemed the only option. Special mention to the evenings of friendship, culture and patatas bravas with my colleague Yovi, which were key for not going crazy during the intense last year of PhD. I also want to thank the work, invisible but essential, by Sonia, who has stood up for us every time a bureaucratic problem appeared. Finally, thanks to the other colleagues from the Statistics and Operations Research Department and the GRBIO research group, especially Lupe, Klaus, and José Antonio, for giving me a hand whenever I needed it.

I also want to thank my friends from Madrid, and those who I left behind in Stuttgart and Liverpool. Luckily, technology brings us closer, and the many reunions we have had make me hope that we will keep what we have built. Thank you also, Josep, for having understood my circumstances when I arrived in Barcelona and for having supported me every time I needed it.

Lastly, thanks to my family. Especially thanks to my parents, my sister and my grandparents, who have trusted me and have given me love and support not only during these years, but throughout my whole life. Without them, this story would never have been written.

Abstract (English)

The lack of transparency and accuracy of research reports has been pointed out as one of the main factors causing research waste. Reporting guidelines (RGs) are sets of recommendations for authors on how to report research methods and findings in a way that no relevant information is missing. Nowadays, there exist more than 400 RGs for different study types, data, and clinical areas. However, biomedical authors' adherence to RGs has been shown to be poor. For this reason, it is warranted to explore what strategies to improve adherence to RGs can be implemented at different points in the research process.

This thesis has three objectives: (i) to identify, classify, and analyse interventions to improve adherence to RGs that have been described in the biomedical literature, and to determine the existing gaps in research on the evaluation of interventions, (ii) to explore biomedical editors' perceptions of different interventions that have been or can be implemented at various points in the editorial process (iii) to evaluate in a real editorial context the impact of an intervention designed based on the studies that address objectives (i) and (ii).

For the first objective, we performed a scoping review of interventions to improve adherence to RGs and identified 31 interventions (11 evaluated, 20 non-evaluated). These were grouped into five categories: training on the use of RGs; improving understanding; encouraging adherence; checking adherence and providing feedback; and involvement of experts. Research gaps identified included the evaluation of interventions (i) on training on the use of RGs and improving understanding of these, (ii) at early stages of research (education, grant writing or protocol writing), and (iii) after

the final acceptance of the manuscript (copyediting or post-publication peer review). Furthermore, we showed that one of the most widespread editorial interventions, the requirement for authors to submit a completed RG checklist together with their manuscript, does not guarantee adherence.

To address the second goal, we performed a survey for biomedical journal editors with experience and interest in the topic of improving authors' adherence to RGs. These editors generally believed that engaging trained professionals in the process of checking adherence to RGs would be the most effective, yet moderately resource intensive, editorial intervention. Also, they thought that standard peer reviewers should not be asked to check RG requirements as they generally lack time and training on the content of RGs. For other promising interventions that could potentially be implemented and evaluated in biomedical journals, we also identified their barriers and facilitators, as well as different types of incentives to encourage the use of RGs.

For our third goal, we carried out a randomised controlled trial. Our goal was to analyse, in a sample of 24 trials submitted to the medical journal BMJ Open, the effect of involving a CONSORT expert in the process of evaluating the submitted checklist and providing feedback to authors. Our results showed that the manuscripts that received this intervention were more completely reported than the ones following the standard process. Based on this, we propose that journals consider revising their peer review processes in order to find ways to make this intervention workable.

In this thesis, we have shown the effectiveness of engaging a reporting expert in the editorial process of a biomedical journal, and we have identified and explored in detail various interventions that future research may consider evaluating. Developing and

implementing effective solutions to improve adherence to RGs is a key step to increase the societal impact of biomedical research and reduce research waste.

Resumen (español)

La falta de transparencia y precisión de los informes de investigación es uno de los principales factores asociados al derroche de recursos financieros invertidos en investigación. Las guías de publicación ("Reporting Guidelines", RGs) especifican cómo los investigadores han de informar de los métodos y resultados de sus estudios, de tal forma que los manuscritos contengan toda la información esencial para los lectores. Hoy en día, existen más de 400 RGs para distintos tipos de estudios, datos y áreas clínicas. Sin embargo, el nivel de adherencia a las RGs es deficiente. Por tanto, es necesario explorar qué estrategias para mejorar la adherencia a las RGs se pueden implementar en distintos momentos del proceso de investigación.

Esta tesis tiene tres objetivos: (i) identificar, clasificar y analizar qué intervenciones para mejorar la adherencia a las RGs han sido descritas en la literatura biomédica, y determinar qué lagunas existen en la evaluación de intervenciones, (ii) explorar las percepciones de los editores biomédicos expertos sobre distintas intervenciones que afectan a los procesos editoriales, y (iii) evaluar el impacto de una intervención diseñada a partir de los estudios relativos a los objetivos (i) y (ii).

Para alcanzar el primer objetivo, realizamos una revisión exploratoria. Esta revisión nos permitió identificar 31 intervenciones que agrupamos en cinco categorías: formación en el uso de RGs; mejora de la comprensión de las RGs; verificación de la adherencia a las RGs y propuestas de mejora para los autores; y colaboración de expertos. Además, identificamos algunas lagunas en las evaluaciones de intervenciones (i) relativas a la formación y mejora de la comprensión de las RGs, (ii) en fases iniciales del proceso de investigación (educación, solicitud de financiación o elaboración de protocolos), y (iii)

después de la aceptación para publicación del manuscrito de investigación (durante el proceso de edición, o la revisión post-publicación del artículo). Por otro lado, mostramos que una de las intervenciones editoriales más populares, que consiste en requerir que los autores completen y envíen la lista de verificación de la RG adecuada junto con su manuscrito, no garantiza la adherencia a esta RG.

En relación con el segundo objetivo, realizamos una encuesta para editores expertos de revistas biomédicas. Estos expresaron mayoritariamente que la intervención potencialmente más efectiva sería involucrar a profesionales formados en el contenido de las RGs, aunque podría requerir un gran volumen de recursos. Además, los participantes apuntaron que los revisores por pares no deberían encargarse de verificar la adherencia a las RGs ya que normalmente carecen de la formación y el tiempo necesarios para realizar esta labor. Finalmente, identificamos las ventajas e inconvenientes de diversas intervenciones prometedoras, así como distintos tipos de incentivos para promover el uso de las RGs.

De cara al tercer objetivo, llevamos a cabo un ensayo aleatorizado con el propósito de analizar, en 24 ensayos aleatorizados recibidos por la revista médica BMJ Open, el efecto de involucrar en el proceso editorial a un experto en CONSORT (la RG para ensayos aleatorizados) que evaluase las guías de verificación enviadas por los autores y les propusiese mejoras. Los resultados señalaron que los manuscritos que pasaban por este proceso eran más completos que los que seguían el proceso estándar. A raíz de esto, proponemos que las revistas ajusten sus procesos de revisión y busquen formas de hacer viable esta intervención.

En esta tesis, hemos demostrado la eficacia de la inclusión en los procesos editoriales de expertos en la presentación de informes científicos. Además, hemos analizado diversas intervenciones que pueden ser evaluadas en el futuro. Desarrollar soluciones efectivas para mejorar la adherencia a las RGs es clave para aumentar el impacto social de la investigación biomédica y reducir el derroche de recursos financieros.

Resum (català)

La manca de transparència y precisió dels informes d'investigació és un dels principals factors associats al malbaratament de recursos financers invertits en investigació. Les guies de publicació ("Reporting Guidelines", RGs) especifiquen com els investigadors han d'informar dels mètodes i resultats dels seus estudis, de manera que els manuscrits continguin tota la informació essencial per als lectors. Avui dia, n'hi ha més de 400 RGs per a diferents tipus d'estudis, dades i àrees clíniques. Tanmateix, el nivell d'adherència a les RGs és deficient. Per tant, és necessari explorar quines estratègies per millorar l'adherència a les RGs es poden implementar en diferents moments del procés d'investigació.

Aquesta tesi té tres objectius: (i) identificar, classificar i analitzar quines intervencions han estat descrites per millorar l'adherència a les RGs en la literatura biomèdica, i determinar quines mancances existeixen en l'avaluació d'intervencions, (ii) explorar les percepcions dels editors biomèdics experts sobre diferents intervencions que afecten als processos editorials, i (iii) avaluar, en un context editorial real, l'impacte d'una intervenció dissenyada a partir dels estudis relatius als objectius (i) i (ii).

Per assolir el primer objectiu, vam realitzar una revisió exploratòria. Aquesta revisió ens va permetre identificar 31 intervencions que vam agrupar en cinc categories: formació en l'ús de RGs; millora de la comprensió de les RGs; verificació de l'adherència a les RGs i propostes de millora per als autors; i col·laboració d'experts. Encara més, vam detectar mancances en l'avaluació d'intervencions (i) relatives a la formació i millora de la comprensió de les RGs, (ii) en fases inicials del procés de recerca (formació, sol·licitud de finançament o elaboració de protocols), i (iii) després de l'acceptació per a publicació

dels manuscrits de recerca (durant el procés d'edició, o la revisió post-publicació de l'article). D'altra banda, vam demostrar que una de les intervencions editorials més populars, que consisteix en requerir que els autors completin i enviïn la llista de verificació de la RG adequada amb el seu manuscrit, no garanteix l'adherència a aquesta RG.

Per al segon objectiu, vam efectuar una enquesta dirigida a editors experts de revistes biomèdiques. Una majoria dels editors van expressar que la intervenció potencialment més efectiva seria involucrar professionals formats en el contingut de les RGs, encara que això podria requerir un gran volum de recursos. Així mateix, els participants van opinar que els revisors per parells no haurien d'encarregar-se de verificar l'adherència a les RGs ja que normalment no tenen el temps i la formació necessaris per realitzar aquesta tasca. Així mateix, vam identificar els avantatges i inconvenients de diverses intervencions prometedores, així com diferents tipus d'incentius per promoure l'ús de les RGs.

En relació amb el tercer objectiu, vam portar a terme un assaig aleatoritzat amb la finalitat d'analitzar, en 24 assaigs aleatoritzats rebuts per la revista mèdica BMJ Open, l'efecte d'involucrar en el procés editorial a un expert en CONSORT (la guia per a assaigs aleatoritzats) que avalués les guies de verificació enviades pels autors i els hi proposés millores. Els resultats van indicar que els manuscrits que passaven per aquest procés eren més complets que els que seguien el procés estàndard. Arran d'això, proposem que les revistes ajustin els seus processos de revisió i busquin formes de fer viable aquesta intervenció.

En aquesta tesi, hem demostrat l'eficàcia de la inclusió en els processos editorials de experts en la presentació d'informes científics. A més, hem analitzat en detall diverses intervencions que poden ser avaluades en el futur. Desenvolupar solucions efectives per millorar l'adherència a les RGs és un pas clau per augmentar l'impacte social de la recerca biomèdica i reduir el malbaratament de recursos financers.

List of contents

List of figures

List of tables

Abbreviations

ARRIVE: Animal Research: Reporting In Vivo Experiments

COBWEB: CONSORT-based web tool

CONSORT: Consolidated Standards of Reporting Trials

EQUATOR: Enhancing the Quality and Transparency of Health Research

E&E: Explanation and Elaboration

ID: Identification

MiRoR: Methods in Research on Research

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

RCT: Randomised Controlled Trial

RGs: Reporting Guidelines

SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

STROBE: Strengthening the Reporting of Observational studies in Epidemiology

Chapter 1: Background and objectives

Approximately 85% of all biomedical research investments is estimated to be wasted (1). Four main problems occurring at different stages of research have been pointed out to explain this alarming number: the choice of wrong research questions; the poor quality of research design and methods; the failure to publish relevant research promptly, or at all; and the bias or lack of usability of research reports. Reducing these problems represents a major societal challenge that has a direct impact on patient care (1).

In this thesis, we aim to focus on the problem of inadequate reporting mentioned above and explore strategies to make research reports more transparent and accurate. The overarching goal of this project is to increase the value of published literature and therefore contribute to the reduction of research waste.

1.1. Meta-research and the MiRoR Project

Meta-research or "research on research" is a field that aims to help science yield better and more reliable results by conducting research on research itself (2). Despite the importance of this discipline, in 2015 only a few initiatives existed all over the world that were performing meta-research and promoting research practices that could improve the efficiency and credibility of scientific investigation. Two prominent examples were the Meta-Research Innovation Center at Stanford (METRICS) and the Centre for Journalology of the Ottawa Hospital Research Institute (OHRI). The lack of similar initiatives in Europe motivated the creation in 2016 of the Methods in Research on Research (MiRoR) Project (3), a joint doctoral training programme funded by Marie

Skłodowska-Curie Actions (grant agreement No 676207) whose aim was to increase research value and reduce research waste. We are fifteen PhD students, seven academic beneficiaries, six non-academic partners, and three academic partners[1]. The different PhD projects in MiRoR cover various areas of meta-research: research methods, research conduct, research reporting, and research evaluation. The present PhD thesis falls within the latter two: the reporting of research results and findings, and the process of evaluation of research evidence or peer review.

In recent years, further initiatives in the field of meta-research have been developed in Europe, such as the Quality, Ethics, Open Science, Translation Center (QUEST) of the Berlin Institute of Health (BIH), the Meta-Research Innovation Center Berlin (METRIC-Berlin), the Meta-Research Center at Tilburg University, and the Research on Research Institute (RoRI).

1.2. Reporting guidelines: promoting transparent reporting of research

As mentioned above, inadequate reporting of research is one of the main factors causing research waste. Transparent and accurate reporting allows researchers to replicate the studies, generate new hypothesis or compare the results of different studies; it allows health care professionals to make clinical decisions; it allows governments to change

[1] **Academic beneficiaries**: Université de Paris, University of Amsterdam, Universitat Politècnica de Catalunya, University of Ghent, University of Split, Centre National de la Recherche Scientifique, and University of Liverpool.
Non-academic partners: The Cochrane Collaboration, The British Medical Journal (The BMJ), BioMed Central (BMC), European Clinical Research Infrastructure Network (ECRIN), National Institute for Health and Care Excellence (NICE), and Sideview.
Academic partners: EQUATOR Network, The Meta-Research Innovation Center at Standford (METRICS), and Ottawa Hospital Research Institute (OHRI).

public policies; and it helps patients to be aware of what healthcare options they have (4).

Reporting guidelines (RGs) are sets of minimum recommendations for authors, usually in the form of a checklist, on how to report research methods and findings so that no relevant information is omitted (4). RGs often consist of two documents (see Figure 1): one that displays the checklist and explains the procedure followed to create it (usually a consensus process among experts in different fields), and another one, the Explanation and Elaboration (E&E) document, that includes a detailed explanation of what authors are expected to report for each of the items in the checklist, along with illustrative examples of adequate reporting.

Table 1 | CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item
Title and abstract		
	1a	Identification as a randomised trial in the title
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts[45 65])
Introduction		
Background and objectives	2a	Scientific background and explanation of rationale
	2b	Specific objectives or hypotheses
Methods		
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons
Participants	4a	Eligibility criteria for participants
	4b	Settings and locations where the data were collected

(a)

Checklist items

Title and abstract

Item 1a. Identification as a randomised trial in the title.

Example—"Smoking reduction with oral nicotine inhalers: double blind, randomised clinical trial of efficacy and safety."[63]

Explanation—The ability to identify a report of a randomised trial in an electronic database depends to a large extent on how it was indexed. Indexers may not classify a report as a randomised trial if the authors do not explicitly report this information.[64] To help ensure that a study is appropriately indexed and easily identified, authors should use the word "randomised" in the title to indicate that the participants were randomly assigned to their comparison groups.

(b)

Figure 1: Example of checklist and items explanation. *Image (a) shows the initial part of the CONSORT checklist. Image (b) contains an example and the explanation for CONSORT Item 1a that can be found in the CONSORT E&E document. These images are reproduced from Moher et al, 2010, (145).*

Since the inception in 1996 of the Consolidated Standards of Reporting Trials (CONSORT) for the reporting of randomised controlled trials (RCTs) (5), more than 400 RGs for different study types, data, preclinical and clinical areas have been developed (6). These can be found in the Library for Health Research Reporting of the Enhancing the QUAlity and Transparency Of Health Research (EQUATOR) Network (6). Apart from CONSORT,

there are some other general RGs for the main study designs, such as STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) for the reporting of observational studies (7) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) for the reporting of systematic reviews (8). There also exist extensions of these RGs that apply to specific study types within a main study design or clinical area. For example, there currently are 17 official extensions of CONSORT (9) that cover different several variations to the standard RCT methodology, including different design aspects (e.g. cluster, pilot and pragmatic trials), interventions (e.g. herbals, acupuncture and non-pharmacologic), and data (e.g. harms and abstracts). Often, more than one RG applies for a certain study. For example, authors of a pilot cluster RCT should use the main CONSORT checklist, as well as the CONSORT extensions for pilot (10) and cluster (11) trials. As it is often challenging for authors to determine what RG(s) apply to their study, an online tool has been developed to help them in this process (12).

RGs are expected to have a central position in the research process: they do not only aim to help authors write their research papers, but also help them at earlier stages of research. If authors are aware of what information will need to be included in their final research reports, they will be more likely to carry out their studies with those requirements in mind and collect the necessary information. Additionally, RGs can also be used by the different stakeholders of the publication process, such as peer reviewers, journal editors or administrators, to check whether the research report is complete enough to be usable by readers.

Adherence to RGs and the methodological quality of a research report are two different concepts that should not be confused. Transparent and accurate reporting is essential

to help readers judge whether the study design and analysis and interpretation were sound. However, adequate reporting is not a measure of the methodological quality of the study as RGs do not intend to dictate the standards for the conduct of research and it should not be used as such (13).

The vast majority of RGs have not yet been assessed as to whether they actually help improve the completeness of reporting of published articles (14) but some, such as CONSORT, have been shown to enhance it (15,16). Dozens of systematic reviews have explored in recent years the extent of author adherence to some RGs in different areas of biomedical research. Samaan et al. (17) went one step further and performed a systematic review that summarised the results of these systematic reviews assessing adherence to RGs. Since they considered a broad range of clinical areas and study designs, their results provided a global picture of adherence to RGs in biomedical research. The authors claimed that, although some studies reported acceptable overall levels of completeness of reporting and found that it had improved since the introduction of certain RGs such as CONSORT, most of the reviews (43 of 50, 86%) concluded that more improvement is needed or that adherence to RGs was inadequate, poor, medium or suboptimal (17). Box 1 includes the definitions of several relevant concepts that we will use throughout the entire document.

Adherence: Action(s) taken by authors to ensure that a research report is compliant with the items recommended by the appropriate/relevant RG. These can take place before or after the first version of the manuscript is published.

Endorsement: Action(s) taken by journals to indicate their support for the use of one or more RG(s) by authors submitting research reports for consideration.

Implementation: Action(s) taken by journals to ensure that authors adhere to an endorsed RG and that therefore published papers are completely reported.

Complete reporting: Pertains to the state of reporting of a study report and whether it is compliant with all the items recommended by the appropriate/relevant RG.

Box 1: Relevant definitions (adapted from Stevens et al, 2014, (14))

1.3. Interventions to improve adherence to RGs

In an attempt to improve the current levels of adherence to RGs of research reports, different interventions have been proposed. The effectiveness of some of these at improving adherence to RGs has been evaluated. For example, biomedical journals have followed strategies ranging from (i) making available editorial statements that endorse certain RGs, (ii) recommending or requiring authors to follow RGs in the "Instructions to authors", and (iii) requiring authors to submit a completed RG checklist together with the manuscript (14–16). Other strategies have been implemented and assessed, such as implementing writing aid tools for authors (18) or involving statisticians in the peer review process (19). While some of these actions have not been shown to have a benefit (14) others report better but still suboptimal levels of reporting (15,16,18,19).

Given the low levels of completeness of reporting that have been observed in the literature (17), it is warranted to explore further interventions to improve adherence to RGs. Furthermore, it is essential to evaluate these in order to provide all the stakeholders of the research process with empirical evidence of the effectiveness of these interventions.

1.4. Thesis objectives and structure

This PhD thesis has three main objectives:

1. To identify, classify, and analyse interventions to improve adherence to RGs that have been described in the published or grey literature, and to determine the existing gaps in research on the evaluation of interventions to improve adherence to RGs (Chapter 2).
2. To explore biomedical editors' perceptions of different editorial interventions that have been or can be implemented at various points in the editorial process (Chapter 4).
3. To evaluate in a real editorial context the impact of an intervention designed based on the results of the previous projects (Chapter 5).

The scoping review described in Chapter 2 is complemented by Chapter 3, which reports the results of a study aimed at exploring one of the most widespread interventions performed by biomedical journals: the requirement for authors to submit a completed RG checklist with their manuscript. To do that, we analyse the degree of consistency between the submitted checklists and the information reported in the manuscripts.

These two initial studies make room for Chapters 4 and 5. In the survey described in Chapter 5, we look into some particularly interesting interventions identified in the initial scoping review by collecting biomedical journal editors' views on issues related to the implementation of these interventions. This survey aims to inform future evaluations of interventions to improve adherence to RGs.

Then, Chapter 5 reports the results of an RCT that we designed based on the results of our previous studies. In view of the poor results of biomedical journals requiring the submission of RG checklists, we aimed to evaluate the effect of engaging a reporting expert who assesses the submitted checklists and provides feedback to authors.

Finally, Chapter 6 reviews the key results of the thesis and provides ideas for further research.

As one of main goals was to explore how the use of RGs can help improve the journal peer review process and the quality of published research reports, all our research except the initial scoping review (Chapter 2) was focused on editorial interventions. These are the interventions that are related to the editorial process of biomedical journals. Moreover, our partner institution, the academic publisher BMJ Publishing Group, gave us the chance to carry out an experimental study (Chapter 5) in collaboration with one of their journals, BMJ Open.

Chapters 2, 3, 4, and 5 are based published papers. In each of these chapters, we have made minor formatting and content edits on the originally published reports. More specifically, we have removed the "Background" sections of the papers, which would be redundant in the context of this dissertation, and focused on the study objectives. For

the other sections (Methods, Results, Discussion), we have made some format changes such as including information that was originally published as supplementary material. In addition, we have corrected some typos and made a few clarifications in light of the comments provided by the thesis reviewers.

As mentioned later in different parts of this document, the supplementary data for each of the projects is publicly available in Zenodo repository (20–23).

Chapter 2: A scoping review to identify and classify interventions to improve adherence to RGs

This chapter is based on the following published research paper (see the last paragraph of section 1.4. to know more about the status of this and the next chapters relative to the published papers):

- **Title:** Scoping review on interventions to improve adherence to reporting guidelines in health research (24)
- **Published in:** BMJ Open, May 2019
- **DOI**: 10.1136/bmjopen-2018-026589
- **PubMed ID:** 31076472
- **Authors:** David Blanco, Doug Altman, David Moher, Isabelle Boutron, Jamie J Kirkham, Erik Cobo

We also published the study protocol for this project:

- **Title:** Interventions to improve adherence to reporting guidelines in health research: a scoping review protocol (25)
- **Published in:** BMJ Open, November 2017
- **DOI**: 10.1136/bmjopen-2017-017551
- **PubMed ID:** 29150467
- **Authors:** David Blanco, Jamie J Kirkham, Douglas G Altman, David Moher, Isabelle Boutron, Erik Cobo

2.1. Study objectives

In this scoping review, our goal was to analyse and classify interventions to improve adherence to RGs in order to obtain a wide picture of how the problem of improving the completeness of research reports had been tackled. The results of this study were expected to help the elaboration of a survey that aimed to look deeper into some of the interventions identified (Chapter 4) and to help design a RCT that would evaluate the effect of one of these interventions (Chapter 5).

More specifically, the research questions for this study were:

1. What interventions to improve adherence to RGs have been evaluated?
2. What further interventions have been performed or suggested but never evaluated?

By answering these questions, we could analyse the implementation details and the effect of interventions that had already been evaluated, as well as to gather other possible strategies that could be implemented and evaluated in the future.

2.2. Methods

We used established scoping review methodology and followed the manual published by the Joanna Briggs Institute for scoping reviews (26). Since we aimed to provide a wide overview of this field, map the key concepts underpinning this research area and the main sources and types of evidence available, we considered that performing a scoping review was the most suitable approach (27).

Eligibility criteria

We included:

1. Studies evaluating interventions aiming to improve adherence to RGs in health research, irrespective of study design.
2. Commentaries, editorials, letters, studies, and online sources describing possible interventions that have been performed or suggested but never evaluated.

We considered the RGs shown on 8 May 2017 on the EQUATOR Network Library for health research reporting (6) as "Reporting Guidelines for main study types". In addition, we included QUOROM (Quality of Reporting of Meta-analyses), since it was the precursor of PRISMA. Table 1 shows all RGs included. We considered the following languages: English, Spanish, French, German, and Catalan.

Exclusion criteria

We excluded references that included interventions that did not specifically aim to improve the completeness of reporting, even though these interventions may actually influence it. For example, we excluded clinical trial registration as a possible intervention (even though it may enhance completeness of reporting), because it mainly aims to reduce publication and selective reporting biases.

Acronym	Full name
CONSORT	Consolidated Standards of Reporting Trials
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SRQR	Standards for Reporting Qualitative Research
COREQ	Consolidated criteria for Reporting Qualitative research
STARD	Standards for Reporting of Diagnostic Accuracy
TRIPOD	Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis
SQUIRE	Standards for Quality Improvement Reporting Excellence
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
SPIRIT	Standard Protocol Items: Recommendations for Interventional Trials
PRISMA-P	Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
CARE	Case Report
AGREE	Appraisal of Guidelines, Research and Evaluation
ARRIVE	Animal Research: Reporting In Vivo Experiments
RIGHT	Reporting Tool for Practice Guidelines in Health Care
QUOROM	Quality of Reporting of Meta-analyses

Table 1: Description of the acronyms and full names of all RGs considered

Search strategy and study selection

On 8 May 2017, we searched PubMed, EMBASE, and Cochrane Library databases for articles published between 1 January 1996 and 31 March 2017, in accordance with our scheduled search (25). Table 2 shows the search terms for PubMed. The search strategy for the other databases can be found in the study protocol (25).

Steps	Search terms
S1	impact* [tw]
S2	improv* [tw]
S3	enhanc* [tw]
S4	boost* [tw]
S5	increas* [tw]
S6	influenc* [tw]
S7	effect [tw]
S8	S1 **OR** S2 **OR** S3 **OR** S4 **OR** S5 **OR** S6 **OR** S7
S9	compliance [tw]
S10	adherence [tw]
S11	completeness [tw]
S12	quality of reporting [tw]

S13	reporting quality [tw]
S14	S9 **OR** S10 **OR** S11 **OR** S12 **OR** S 13
S15	Consolidated [tw] Standards [tw] Reporting [tw] Trials [tw] **OR** CONSORT[tw]
S16	Strengthening [tw] Reporting [tw] Observational [tw] Studies [tw] Epidemiology[tw] **OR** STROBE[tw]
S17	Preferred [tw] Reporting [tw] Items [tw] Systematic [tw] reviews [tw] Meta-Analyses [tw] **OR** PRISMA[tw]
S18	Standards [tw] Reporting [tw] Qualitative Research[tw] **OR** SRQR[tw]
S19	Consolidated [tw] Criteria [tw] Reporting [tw] Qualitative [tw] Research[tw] **OR** COREQ[tw]
S20	Standard [tw] Protocol [tw] Items [tw] Recommendations [tw] Interventional [tw] Trials[tw] **OR** STARD[tw]
S21	Transparent [tw] Reporting [tw] multivariable [tw] prediction [tw] model [tw] Individual [tw] Prognosis [tw] Diagnosis[tw] **OR** TRIPOD[tw]
S22	Standards [tw] QUality [tw] Improvement [tw] Reporting [tw] Excellence[tw] **OR** SQUIRE[tw]
S23	Consolidated [tw] Health [tw] Economic [tw] Evaluation [tw] Reporting [tw] Standards[tw] **OR** CHEERS[tw]
S24	Standard [tw] Protocol [tw] Items [tw] Recommendations [tw] Interventional [tw] Trials[tw] **OR** SPIRIT[tw]
S25	Preferred [tw] Reporting [tw] Items [tw] Systematic [tw] Review [tw] Meta-Analysis [tw] Protocols[tw] **OR** PRISMA-P[tw]
S26	Quality [tw] Reporting [tw] Meta-analyses[tw] **OR** QUOROM[tw]
S27	Case [tw] Report [tw] **AND** CARE[tw]

S28	Appraisal [tw] Guidelines [tw] Research [tw] Evaluation[tw] **AND** AGREE[tw]
S29	Animal [tw] Research [tw] Reporting [tw] Vivo [tw] Experiments[tw] **AND** ARRIVE[tw]
S30	Reporting [tw] Tool [tw] Practice [tw] Guidelines [tw] Health [tw] Care[tw] **AND** RIGHT[tw]
S31	S15 **OR** S16 **OR** S17 **OR** S18 **OR** S19 **OR** S20 **OR** S21 **OR** S22 **OR** S23 **OR** S24 **OR** S25 **OR** S26 **OR** S27 **OR** S28 **OR** S29 **OR** S30
S32	S8 **AND** S14 **AND** S31
S33	S32 **AND** "1996/01/01"[PDAT] : "2017/03/31"[PDAT]

Table 2: Search terms for MEDLINE (via PubMed)

We exported the retrieved studies into Mendeley reference manager (28) and automatically removed the duplicates using it. First, one reviewer (DB) screened the titles and abstracts for eligibility. Each of the other two reviewers (JJK and EC) was randomly assigned 50% of the references and screened the titles and abstracts independently of the first reviewer. The reviewers classified the references into one of the following groups:

A) Evaluated: Includes references describing interventions to improve adherence to RGs that have been empirically assessed.
B) Non-evaluated: Includes references describing interventions to improve adherence to RGs that have been performed or suggested but never evaluated.

C) Unclear: Includes references (i) containing vague statements such as “Authors, editors, and journals have to adhere better to RGs to improve the quality of reporting” or “greater efforts have to be made by authors to check that their research is compliant with [the relevant RG]”, or (ii) not having the abstract available.

D) Excluded: Includes references (i) not describing interventions to improve adherence to any of the RGs considered and (ii) describing but not evaluating certain interventions that have already been classified as evaluated.

Disagreements between reviewers were solved by consensus.

Second, one reviewer (DB) examined the full-text of all group A and B references to confirm the previous classification, then all group C references to reclassify them either as group A, B, or D. The initial reviewer (JJK or EC) verified the reclassification. Finally, one reviewer (DB) ensured literature saturation by searching the reference lists of included studies, the lists of articles citing them according to PubMed, and the individual studies included in two relevant systematic reviews (14,16).

In addition, we performed a grey literature search, which included: the websites of networks and organizations promoting the use of RGs (i.e. EQUATOR Network and National Library of Medicine Research Reporting Guidelines and Initiatives); work groups of medical journal editors (i.e. International Committee of Medical Journal Editors (ICMJE) and World Association of Medical Editors (WAME)); biomedical journal publishers (i.e. BMJ Publishing Group and BioMed Central); funding agencies (i.e. National Institute of Health (NIH) and European Research Council); online platforms of post-publication peer review (i.e. PubPeer and ScienceOpen); and the abstract books of

the past editions of the International Congress on Peer Review and Biomedical Publication.

Some of the included references were described in studies co-authored by some of the authors in this scoping review. These references underwent the same process of screening, data extraction, and data synthesis as the others.

Data extraction

We developed a data extraction form to collect the information necessary for data synthesis. Two reviewers (DB, JJK) independently performed a pilot data extraction on a random sample of five articles and subsequently refined the form.

Extracted data included:

1. Publication characteristics: title, year of publication, author, author's affiliation country, and field of study.
2. Characteristics of the intervention:
    a. Classification as evaluated or non-evaluated.
    b. Research stage: education, grant writing, protocol writing, manuscript writing, submission, journal peer review, copy-editing, and post-publication.
    c. Rationale of the intervention, which refers to the deduced reasons why the intervention is evaluated or proposed.
    d. For evaluated interventions: details of the intervention, study design (e.g. RCT, before-after, etc.), RGs considered and format (checklist, bullet points and/or examples), period of intervention, number of journals and

articles involved, effect size of the intervention on adherence to RGs and measure used to assess this effect.

3. Relevant conclusions.

Two reviewers (DB, JJK) independently performed data extraction for all studies except for the individual studies of the two systematic reviews evaluating journal endorsement of RGs (14,16), since none of these studies described further interventions and their results had already been reported in these reviews. Disagreements between reviewers were solved by consensus.

Data synthesis

Following data extraction, we categorised interventions as follows:

1. Training on the practical use of RGs: mentoring of different stakeholders on the practical use of RGs.
2. Enhancing accessibility and understanding: dissemination of RGs and the improvement of authors' understanding of their content.
3. Encouraging adherence: suggestions and tools to facilitate compliance.
4. Checking adherence and providing feedback: checking the level of compliance and indicating incorrect or missing items.
5. Involvement of experts: interaction and cooperation on methodology and reporting.

One reviewer (DB) performed the initial categorization. The other two reviewers (JJK and EC) verified it.

Furthermore, we determined the existing gaps in research on the evaluation of interventions to improve adherence to RGs. More specifically, we identified which categories of interventions and which research stages have not been addressed so far in studies evaluating interventions.

We did not perform a meta-analysis of the observational studies assessing journal endorsement of RGs that were not included in the two systematic reviews previously mentioned (14,16). We considered that, for the purpose of this scoping review, these systematic reviews provided a reliable picture of the impact of this editorial intervention.

Deviations from the protocol

In order to better capture the most relevant aspects of the included studies, we modified the original data extraction form proposed in the protocol. We removed the health care area of the studies included, refined the research stages considered, and included more details on the implementation of the evaluated interventions.

2.3. Results

The database search yielded 1399 citations after removing duplicates. Screening of titles and abstracts resulted in a first classification, after which we included 435 papers for full text review. We also reviewed the full text of 24 additional references found through forward citation searching. Furthermore, a grey literature search yielded seven additional references. Finally, we included 109 references. Some of these interventions appeared in more than one reference and some of the references contained more than one intervention. 90 of these references (86 observational and 4 randomised studies)

described 11 evaluated interventions and the other 19 (12 research studies, 2 editorials, blogs, 1 commentary, 1 essay, and 1 perspective) described 20 non-evaluated interventions. Figure 2 shows the flow diagram of the study.

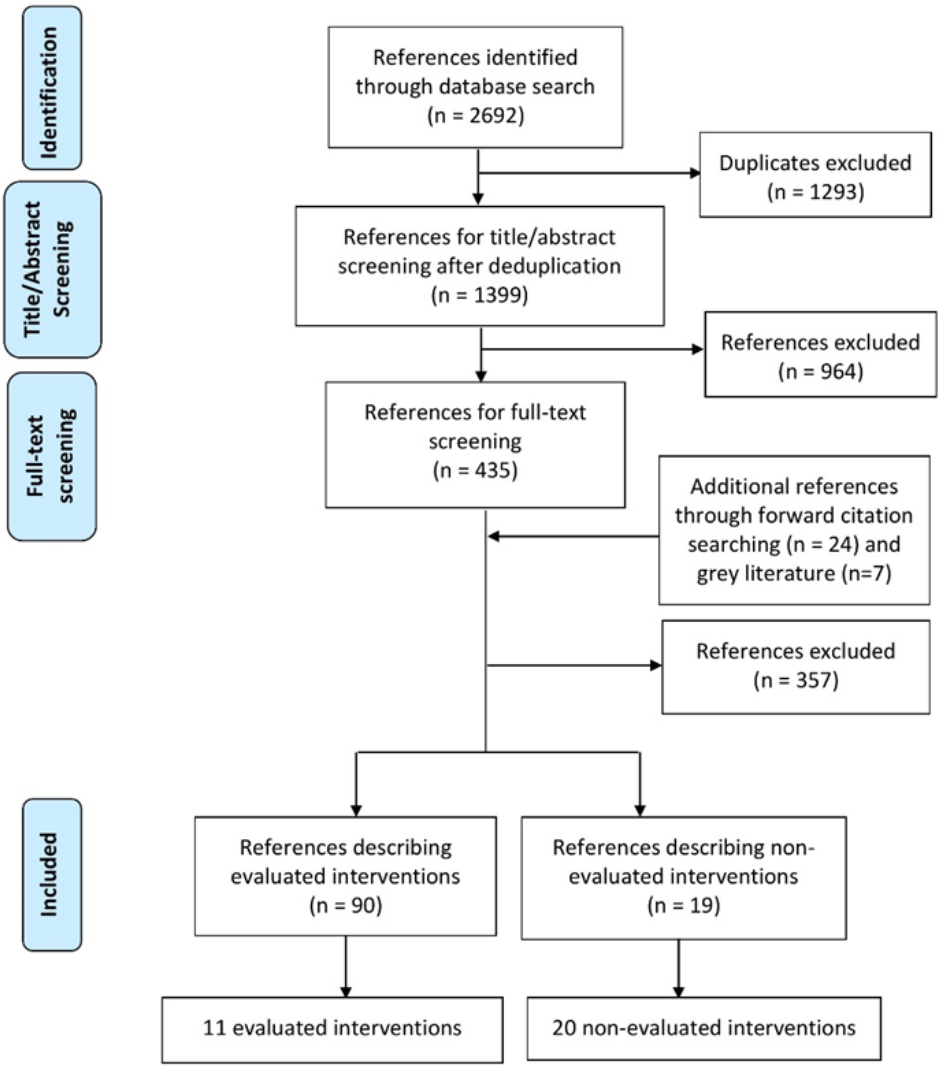


Figure 2: Flow diagram of the scoping review

The 31 interventions identified are displayed in Figure 3 according to their categorization and the research stage where they can be performed. Moreover, Table 3 shows all interventions in a tabular format together with their rationale. The list of all included references include and raw data extracted from them can be found in Zenodo (20).

Group	Intervention	Rationale
Training on the practical use of RGs	Introduction of RGs & journalology into graduate curricula (29–33)	To introduce good research reporting habits early in young researchers' scientific careers.
	Student's development of protocols for coursework and research using RGs (32)	
	Funder's support of author training on RGs (1)	Authors, editors, and peer reviewers have insufficient training in issues related to reporting.
	Training for peer reviewers and editors on RGs by journals (1,33)	
Enhancing accessibility and understanding	Dissemination of RGs by scientific associations (34)	A large number of researchers are not aware of the existence of RGs.
	Translation of RGs to further languages (35)	Language barriers may affect the proper use of RGs.
	Development of expanded database of examples for each RG (36)	Authors need more examples of good reporting to properly understand certain items.
Encouraging adherence	Author use of RGs as a template for grant application proposals (32)	Using RGs in early stages may facilitate completeness of

	Required checklist for ethics approval application (17)	reporting of published research.
	Funder's requirement of checklists in author's report (32,37)	
	Author use of the writing aid tool COBWEB (18)	A) Authors need help to successfully adhere to RGs at the writing stage and B) Dividing RG items into bullet points and providing examples might help.
	Author use of a structured approach for reporting research (38,39)	A) To help authors avoid omissions, B) to aid reviewers and editors in appraising articles and C) to allow more efficient data extraction during the systematic review process.
	Author markup of the manuscript to indicate where each RG item is addressed (40)	
	Editorial statement endorsing certain RGs (41,42,51–60,43,61–70,44,71–80,45,81–90,46,91–100,47,101–110,48,111–120,49,50)	Authors read editorial statements and follow “Instructions to authors”.
	Recommendation or requirement to follow RGs in the "Instructions to authors“ (41,42,51–60,43,61–70,44,71–80,45,81–90,46,91–100,47,101–110,48,111–120,49,50)	
	Requirement to submit a RG checklist together with the manuscript indicating page numbers corresponding to each item (41,42,51–60,43,61–70,44,71–80,45,81–90,46,91–100,47,101–110,48,111–120,49,50)	Authors may not consider editorial statements or recommendations in “Instructions to authors” to be important. Compulsory submission of checklists or text mark-up may encourage

	Requirement to populate and submit a RG checklist with text from the manuscript (121)	authors to be more compliant with RGs.
	Journal development of core versions of RGs containing key items (122)	Focusing on the most important items could be more effective than considering the whole checklist.
	Guidance to authors on manuscript preparation by publication officers (123)	Trained journal officers may enhance authors' compliance with RGs during manuscript preparation.
	Suggestion for peer reviewers to use RGs (124)	Peer reviewers often do not detect reporting flaws. Therefore, they may need to follow a more systematic approach and use RGs.
	Editor's questions to peer reviewers about whether the authors have followed RGs (125)	
Checking adherence and providing feedback	Completeness of reporting check by editors (126)	Requiring checklists at submission does not guarantee adherence. Editors and peer reviewers have to check whether submitted papers are compliant with RGs.
	Peer review against RGs (19)	
	Internal peer review against RGs by a trained editorial assistant (127)	It is extremely unlikely that the average clinical peer reviewer has the methodological expertise to check a paper against RGs.
	Implementation of the automatic tool StatReviewer (128)	
	Email to authors to revise the manuscript according to RGs (129)	It might be more effective to ask authors for adherence to RGs during the revision

	Implementation of the tool WebCONSORT (130)	process because they will do anything to get their paper published.
	Completeness of reporting check at copy-editing (131)	Copy-editing and post-publication offer alternate time points to improve adherence to RGs.
	Post- publication peer review (132)	
Involvement of experts	Statistician involvement (95,133–135)	Professionals with specific knowledge of RGs might help authors when designing, conducting or reporting their research.
	Medical writer involvement (37)	

Table 3: Interventions identified and their rationale

Among the 11 evaluated interventions identified, we found a variety of measures used to assess their effect on adherence to RGs, including:

- Score for completeness of reporting for each paper, either assigning different or equal weights to RG items, on a 0-10 scale.
- Percentage of items reported for each paper.
- Percentage of compliance per RG item.
- Score for the Manuscript Quality Assessment Instrument (MQAI) (136) for each paper.

Appendix A shows the effect sizes of the evaluated interventions, as well as their implementation details.

Four of the 11 evaluated interventions identified were assessed in RCTs. In these trials, statistically significant effect of the intervention was only observed for the use of the writing aid tool for authors COBWEB (18). While performing an additional review against RG showed slightly positive but not significant effect (19), suggesting the use of RG by peer reviewers (124) or implementing at the process of author revision of the manuscript the web-based tool WebCONSORT showed no benefit (130). The rest of the evaluations of interventions found (86 of 90) were observational studies. Most of these evaluated the impact of journal endorsement on completeness of reporting and concluded that it was suboptimal (14,16). However, completeness of reporting improved remarkably when editors got involved in the process of checking adherence to RGs (126) and when research results were posted in a tabular format without discussion or conclusions (38).

Research gaps identified (see Figure 4) included the evaluation of interventions (i) on training on the use of RGs and improving understanding of these, and (ii) at early stages of research (education, grant writing or protocol writing), and (iii) after the final acceptance of the manuscript (copyediting or post-publication peer review).

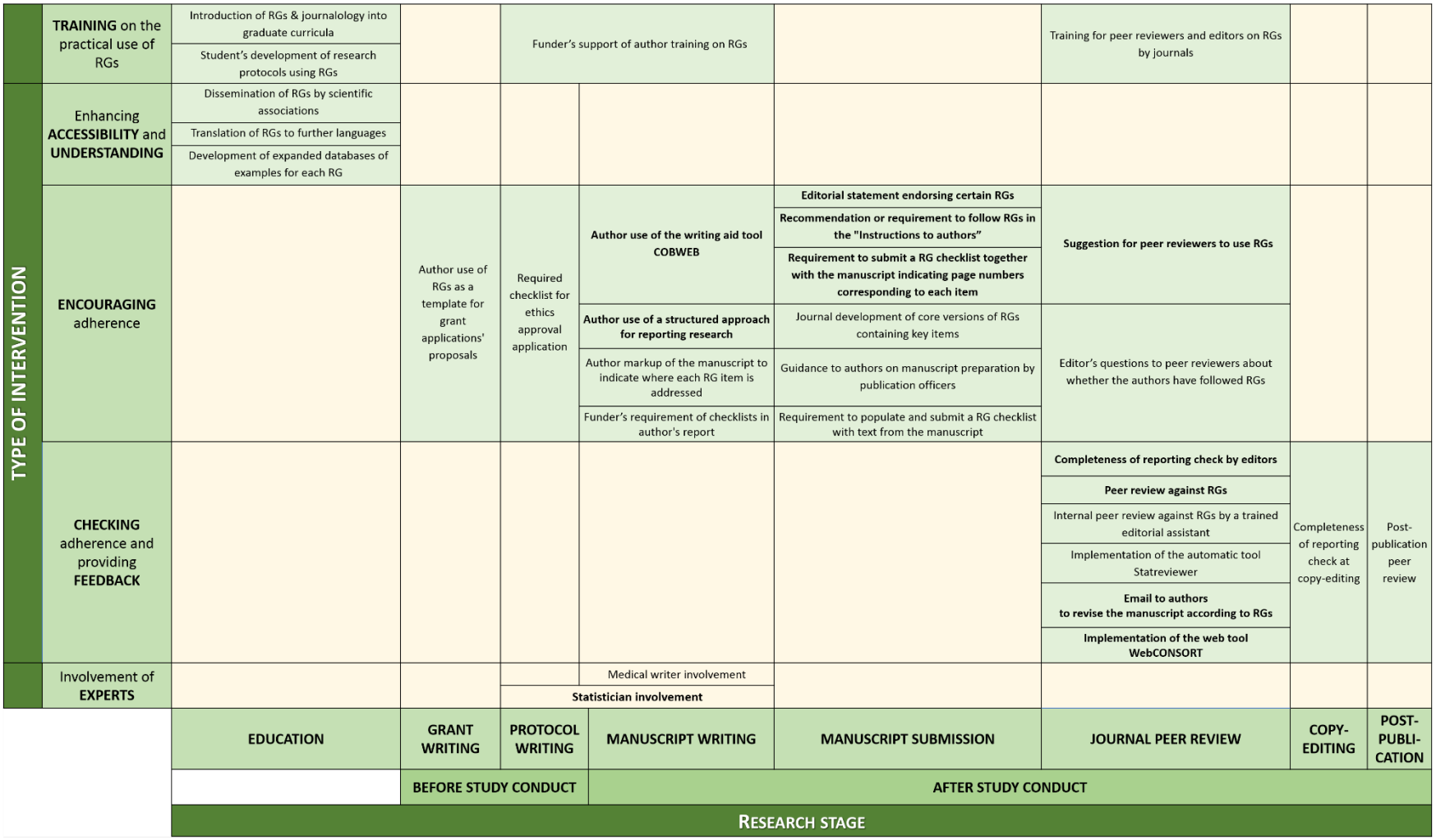


Figure 3: Typology of interventions according to type of intervention and research stage. *Evaluated interventions are shown in bold.*

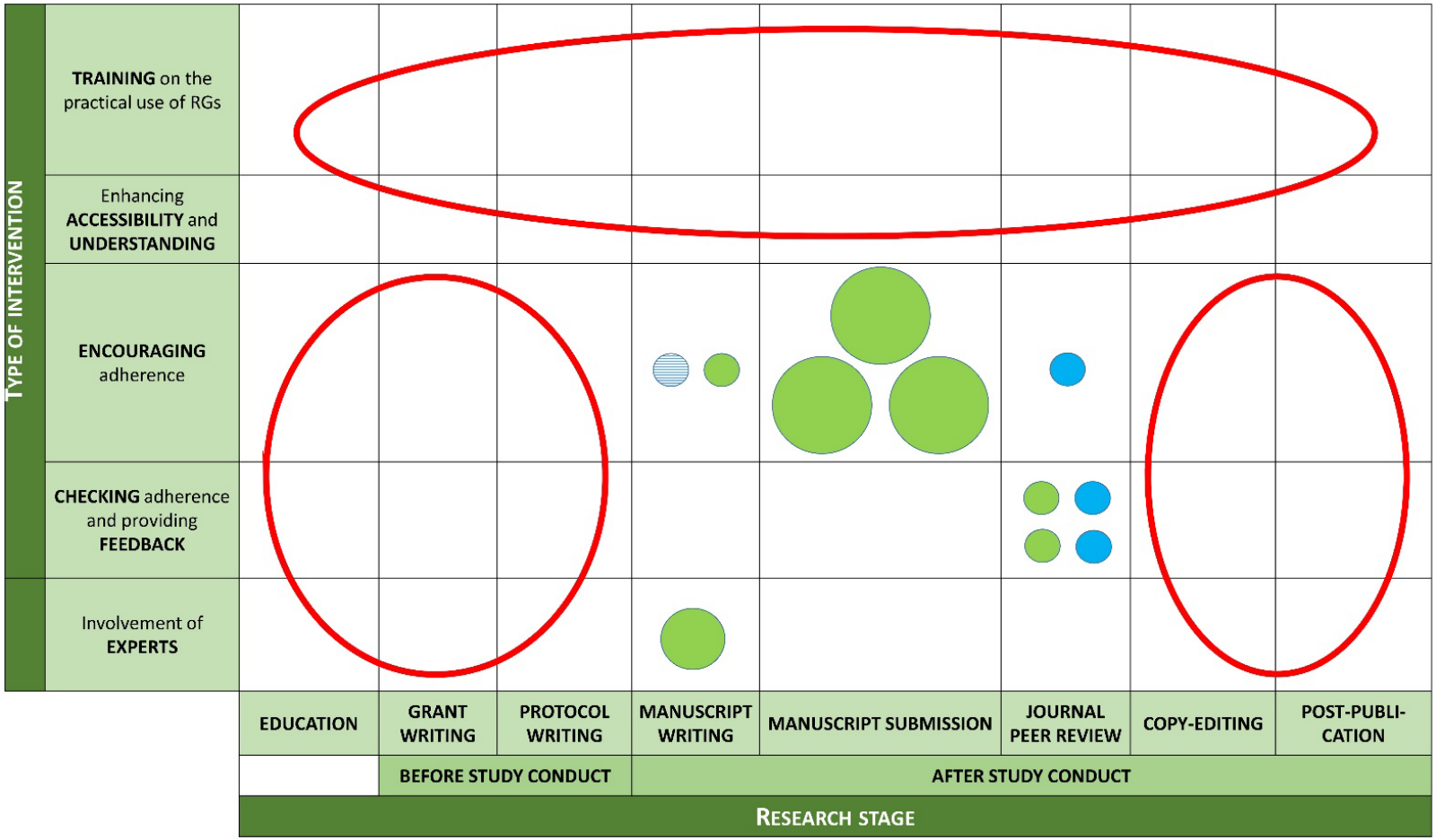


Figure 4: Gaps in research on the evaluation of interventions to improve adherence to RGs. *Each circle represents one intervention. Variables displayed: 1) Circle size: Number of studies evaluating each intervention (bigger = more studies); 2) Circle colour: Study design of those studies (blue for RCTs and green for observational studies) and 3) Circle fill: Kind of RG implementation (plain for checklist and stripes for bullet points and examples). Research gaps are highlighted in red.*

Hereafter, we describe the interventions found for each category. Table 3 and Figure 3 summarise these interventions and Appendix A analyses the details of the evaluated interventions.

Training on the practical use of RGs

Four non-evaluated interventions related to educating different stakeholders on the practical use of RGs were found (18-23).

In a first step, health profession schools could incorporate RGs into curricula that address research methodology and publication standards (29–33). In line with this, students could develop protocols for coursework and research using RGs such as SPIRIT (RCTs) and PRISMA-P (systematic reviews), and educators may encourage adherence to those guidelines and grade the protocols using them (32). For their part, funders may consider supporting author training on RGs (1). Finally, journals or publishers may consider investing resources in training editors and reviewers on the content and use of RGs (1,33).

Enhancing accessibility and understanding

We identified three non-evaluated interventions that were focused on increasing the awareness of the existence of RGs, as well as the authors' understanding of content of these (24-26).

First, international scientific associations may play an important role in disseminating and popularizing RGs to large audiences (34). Second, RG developers might consider translating them to new languages that have not been addressed yet (35). Finally,

further databases of examples of good reporting for different RGs that are accessible to authors can be developed, as has been done for CONSORT (36).

Encouraging adherence

Fourteen interventions found were associated with different strategies to facilitate compliance with RGs (17,18,44–53,32,54–63,37,64–73,38,74–83,39,84–93,40,94–103,41,104–113,42,114–123,43,124,125). Six of these were evaluated (18,38,49–58,41,59–68,42,69–78,43,79–88,44,89–98,45,99–108,46,109–118,47,119,120,124,48).

Funders might require authors to use RGs as a template for grant application proposals (32). Later on, research ethics boards may require that protocols submitted for ethical approval clearly state which RGs the study will be using based on the study design, and that RG checklists are part of the application for ethics approval (17). Funders could also encourage adherence by asking for RG checklists as part of the authors' report (32,37).

One initiative to support authors adhering to RGs at the writing stage of the manuscript has been COBWEB, a writing aid tool that aims to help authors adequately combine the different extensions of the CONSORT statement (18). This tool divided the CONSORT items into bullet points showing the key elements that need to be reported together with examples of adequate reporting. The impact of COBWEB was evaluated in an RCT that showed a large effect of this intervention (18) (see Appendix A for more details about this and other evaluated interventions). A second option to support authors at manuscript writing is that they follow a more structured approach. For example, ClinicalTrials.gov requires authors to report key information in a tabular format when registering a study or making available its results (137). This has been shown to be

effective: some results posted on this platform, especially harms, are more complete than those in corresponding journal articles reporting the same trials (38). Another possibility to improve the structure of manuscripts is to include new subheadings corresponding to different RG items within the traditional IMRaD format (Introduction, Methods, Results, and Discussion), as the American Journal of Orthodontics and Dentofacial Orthopedics (AJO-DO) proposed (39). Finally, authors may also avoid omissions when writing the manuscript if mark up the text and highlight where each item of the relevant checklist is addressed (40).

At manuscript submission stage, different editorial actions have been taken to improve adherence to RGs. The most popular is what has traditionally been defined as journal endorsement of RGs, which is usually defined as one or more of the three following interventions: (a) journal editorial statement endorsing certain RGs; (b) requirement or recommendation in journal's 'Instructions to Authors' to follow certain RGs when preparing their manuscript; or (c) requirement for authors to submit the appropriate RG checklist together with their manuscript indicating page numbers corresponding to each item (16). Dozens of observational studies have explored the possible effect of journal endorsement of different RGs in different clinical areas (41,42,51–60,43,61–70,44,71–80,45,81–90,46,91–100,47,101–110,48,111–120,49,50). A recent systematic review focused on CONSORT evaluations showed relative but suboptimal improvements in the completeness of reporting in journals by following the aforementioned policies (16), while another systematic review considering nine other guidelines showed no improvements (14).

Journals might also consider other strategies to enhance adherence to RGs at submission. A first option could be to develop shorter, core versions of RGs containing key items, which could be provided to authors as part of the submission process (122). Second, they might introduce publication officers in order to provide guidance to authors on preparing manuscripts for submission (123). Third, editors may ask authors to populate the relevant checklist with text from their manuscript and not accept a submission unless this is provided (121).

Finally, editors may suggest that peer reviewers use RGs (124). In addition, by asking peer reviewers questions about whether the author has followed RGs, this might be an indirect way to encourage them (125).

Checking adherence and providing feedback

Eight interventions were related to monitoring level of compliance with RGs of the manuscripts and providing instructions to authors on how to improve the reporting of missing or incorrect items (19,126–132). Four of them were evaluated (19,126,129,130).

Some journals have opted for implementing RGs at peer review. First, an associate editor may assess manuscripts for adherence to the relevant RG and ask authors to make changes accordingly (126). This process may be repeated until the associate editor thinks that the manuscript can move to the next step of the review process, leading to an editorial decision. This intervention was evaluated at the AJO-DO and showed satisfactory results: 33 of 37 items reached perfect compliance (126). Second, peer reviewers could also assess the manuscripts against the appropriate checklist (19). While the observed effect of this intervention was slightly positive, it was smaller than

hypothesised. In fact, investigators pointed out that authors tended to comply better with suggestions coming from standard reviews rather than from reviews against RGs, implying that it might be difficult to adhere to high methodological standards at late stages of research if these standards are not considered earlier in the research process. Third, journals could also ask trained editorial assistants to check manuscripts against RGs (127) or to implement automatic peer review tools such as StatReviewer, software that automatically checks adherence to RGs and evaluates the appropriate use and reporting of statistical tests (128). Currently, its performance is being assessed through a pilot trial in collaboration with four BioMed Central Journals (128). In any of those cases, emails could be sent to authors asking them to revise the manuscript according to guidelines (129). To do this, the EQUATOR Network has provided standard letters that can be used a) after checks by an editor or a single peer reviewer, b) after full peer review, or c) alongside acceptance (138). Furthermore, at the time of author revision of the manuscript, Hopewell et al. found no significant effect when incorporating WebCONSORT, a web-based tool that generates a unique list of items customised to the trial design, to the revision process of journals that endorsed CONSORT but had no active policy for implementing it (130). Finally, in a late stage of the publication process, copyediting of the manuscript could also ensure that all items are covered (131).

Once the paper is published, the scientific community could use online platforms of post-publication peer review such as PubPeer (139) or ScienceOpen (140) to evaluate the adherence to RGs of published articles and to provide feedback to authors (132).

Involvement of experts

Two interventions identified implied interaction and cooperation between authors and experts on methodology and reporting at different stages of research (37,95,133–135). One of them was evaluated (95,133–135).

On the one hand, statisticians (or epidemiologists or other quantitative methodologists) may get involved in the design, conduct or reporting of the study might contribute to properly reporting key areas such as sample size calculation, randomisation, blinding, and appropriate statistical analysis (134). While three studies found a statistically significant positive relationship between CONSORT scores and statistician involvement (95,134,135), another one did not (133). On the other hand, it has been hypothesised that the involvement of medical writers during the manuscript writing stage of research could improve the completeness of reporting (37).

2.4. Discussion

In this scoping review, we identified 31 interventions to improve adherence to RGs. We also determined the gaps in research on the evaluation of this type of interventions. By considering a wide range of RGs as well as their extensions and merging the evidence found in the published and grey literature, this review provides a broad picture of how the problem of enhancing adherence to RGs has been tackled so far and could be faced in the future.

This study reveals that most published research aimed at improving adherence to RGs has been conducted in journals. Typically, journal strategies range from making available editorial statements that endorse certain RGs, recommending or requiring authors to

follow RGs in the “Instructions to authors”, and requiring authors to submit RG checklist together with the manuscript, with page numbers indicated for each item. However, these strategies have been shown not to have the desired effect (14,16). Recent research has called for more active and enforced journal policies throughout the editorial process, such as requiring the use of structured approaches with new subheadings adapted to different kinds of study designs (39), which was also found to be beneficial in a new study outside of our search period (141); providing guidance on manuscript preparation (123); making sure the peer review process involves editorial assistants who have specific training on reporting issues (127); and implementing automatic peer review tools (128). Journals will vary in their ability to make some of these strategies effective, depending on factors such as their resources, their guidelines to peer reviewers and the dedication of their editors – many editors and editorial staff work part-time and have limited amount of time.

Moreover, editors’ education and performance should be improved. A recent study pointed out that more than a third (39%) of the manuscripts classified as RCTs by the editorial staff were not actually RCTs (130,142). Consequently, it seems difficult to improve author adherence to RGs if journal gatekeepers are not properly trained in methodological and reporting issues.

Apart from journals, editors and peer reviewers, other key stakeholders such as medical schools, research funders, universities and other research institutions should also take responsibility regarding this issue. This scoping review provides some strategies to follow. However, as the problem is complex and the possible interventions are varied, enhancing the completeness of reporting most likely depends not so much on any

isolated action but on a set of strategies by several different stakeholders. These could be enacted at different stages of research, from education to article post-publication.

For interventions aiming to improve adherence to RGs, we should require the same level of evidence that we require for interventions to improve health. For this reason, it is striking that we found only four published RCTs that evaluated interventions to improve adherence to RGs (18,19,124,130). Among these trials, statistically significant effect of the intervention was only observed for the use of the writing aid tool for authors COBWEB (18). While performing an additional review against RGs showed slightly positive but not significant effect (19), suggesting the use of RGs to peer reviewers (124) or implementing at the process of author revision of the manuscript the web-based tool WebCONSORT showed no benefit (130). The rest of the evaluations of interventions found (86 of 90) were observational studies, whose results are subject to the influence of confounding factors. As already mentioned, the impact of journal endorsement on completeness of reporting was suboptimal (14,16). However, completeness of reporting improved remarkably when RGs were actively implemented by editors (e.g. if editors perform a completeness of reporting check of the manuscript (126)) and when research results were posted in a tabular format without discussion or conclusions (38). Future RCTs should consider evaluating these interventions or addressing some of the research gaps identified in this review, such as improving adherence to RGs at the grant application or protocol writing stages.

A few of the interventions found in this review were shown to enhance adherence to RGs. However, it is noteworthy there is no evidence that some successful interventions (18,126) have been implemented more widely later. For this reason, more resources and

efforts are needed to further implement these interventions in other settings, evaluate the effect, and share the results with the scientific community. In any case, it is important to keep in mind that contemporary publication culture may harm the potential improvements in reporting quality. This could result from the fact that most scientists feel that the primary evaluation tool of their research is the quantity of their scientific output rather than its quality (143); and such attitudes may undermine the potential effect of any intervention to improve adherence to RGs.

Our scoping review has some limitations. First, we did not formally assess the methodological quality of the studies that evaluated interventions. Second, restricting to certain databases or not having standard search terms for the databases searched may have excluded relevant publications. Third, it is possible that we could have missed evidence of possible interventions that may have never been reflected in the published or grey literature but are instead used in practice and continue to be used. For example, journals might be applying specific editorial strategies that are not publicly available on their websites or in the published literature.

Improving adherence to RGs is one of the key issues in order to enhance complete and accurate reporting and therefore reduce waste in research. Different stakeholders – such as research funders, ethics boards, and journals – should consider implementing and evaluating some of the interventions identified in this study.

Chapter 3: Analysing the effect of a widespread intervention to improve adherence to RGs

This chapter is based on the following publication:

- **Title**: Are CONSORT checklists submitted by authors adequately reflecting what information is actually reported in published papers? (144)
- **Published in:** Trials, January 2018 as a commentary
- **DOI**: 10.1186/s13063-018-2475-0
- **PubMed ID**: 29378640
- **Authors**: David Blanco, Alice M. Biggane, Erik Cobo, MiRoR Network*

*Members of the MiRoR Network involved: Doug Altman, Lorenzo Bertizzolo, Isabelle Boutron, Efstathia Gkioni, Ketevan Glonti, Jamie Kirkham, Camila Olarte, Maria Olsen, Cecilia Superchi.

Since we submitted this contribution to Trials journal as a commentary, it did not follow the traditional IMRaD format (Introduction, Methods, Results, and Discussion). We believed this was the most appropriate format since we aimed to present a short article covering an issue related to the editorial policies of Trials journal (and to many other journals) and therefore especially relevant to its scope.

3.1. Study objectives

As mentioned in the previous chapter, one of the most widespread journal strategies to improve adherence to RGs is to require authors to submit a completed RG checklist (and the relevant RG extension(s)) together with the manuscript. In addition, in an effort to make the editorial process more transparent and credible, some journals following this policy, such as PLOS One, BMJ Open, and Trials, also make the checklists submitted by the authors accessible for the readers as supplementary material.

Our goal was to illustrate whether the submission of completed RG checklists should be considered a guarantee that the RG items are adequately reported. Here, we focused on the CONSORT guidelines and its extensions as CONSORT is currently one of the most well-established RGs and has been revised and updated twice (5,145,146).

Our specific objectives were to explore 1) whether authors were submitting the appropriate CONSORT checklists, and 2) whether the completed checklists submitted by authors were consistent with the information reported in the manuscript.

3.2. Our findings

We searched in PubMed on 12 June 2017 for RCTs published in PLOS One, BMJ Open, and Trials between 1 January 2016 and 6 June 2017 using the search terms shown in Table 4. We chose those journals because they request the submission of completed CONSORT checklists for RCTs and they make these checklists accessible for the readers as supplementary material.

Steps	Search terms
S1	Trials [Journal]
S2	BMJ Open [Journal]
S3	Plos ONE [Journal]
S4	S1 **OR** S2 **OR** S3 **NOT** Protocol
S5	S4 **AND** Randomized Controlled Trial[ptyp] **AND** ("2017/01/01"[PDAT] : "2017/06/06"[PDAT])

Table 4: Search terms for MEDLINE (via PubMed)

The search returned 232 hits (176 from PLOS ONE, 36 from BMJ Open, and 20 from Trials). We used R software (147) to randomly select five papers from each journal. Some of the papers initially selected were not suitable for further analysis because either they were not RCTs or the CONSORT checklist was not available. Therefore, we excluded those papers and randomly selected new ones for PLOS ONE and BMJ Open until we had five for each journal. For Trials, we could only find two papers meeting the inclusion criteria. Therefore, the final analysis included twelve papers.

For each paper included, we retrieved the initial CONSORT checklist and manuscript submitted by the authors. First, we independently determined if the CONSORT checklist originally submitted by the authors was the appropriate extension for the study design. Then, for papers using the appropriate checklist, we compared it with what was actually reported in the published paper to identify any inconsistencies. To do this, we randomly

split the twelve selected papers into three groups of four papers. One of the main investigators (AB, DB, and EC) reviewed each of the groups, as well as two of the six collaborators (LB, EG, KG, CO, MO, and CS) from the MiRoR Network. Prior to this assessment, all evaluators had participated in a training session on the content of CONSORT guidelines and had practised the identification of information related to the core CONSORT items (see the following paragraph). Disagreements among evaluators were solved by consensus.

We focused our evaluation on six core CONSORT items of the "Methods" and "Results" sections: (6a) outcomes; (8a) sequence generation; (9) allocation concealment mechanism; (11a) blinding; (13a) flow of participants; and (13b) losses and exclusions. We selected those items because they are essential for systematic reviewers to evaluate the risk of bias (148) and are known to be poorly reported (127). For each item, we used the CONSORT Explanation and Elaboration (E&E) document (149) to determine what information we expected authors to report.

For each CONSORT item, we graded the consistency between what authors claimed in the checklist and what they reported in the manuscript as follows:

- Completely consistent: There was no divergence between what authors claimed to have reported through the originally submitted CONSORT checklist and what they reported in the published paper.
- Partially consistent: This may include any of the following: (a) Partial absence of relevant information that was expected to be reported; or (b) the information corresponding to that item was reported elsewhere in the paper to what was claimed in the checklist.

- Not consistent: Authors claimed to have reported that item through the CONSORT checklist but did not adequately report the information in the published paper.

From the twelve randomly selected RCTs, the standard CONSORT checklist was appropriate for six papers (four of which were standard parallel trials covered by the standard CONSORT and two were crossover trials, for which there was not an extension at the time we carried out the study). The other six required CONSORT extensions (for cluster trials, three; for pragmatic trials, two; and for non-pharmacological interventions, one) but authors did not use them in any case, despite being available at the time of submission. The aforementioned extensions were published between 2008 and 2012 (11,150,151), yet the six papers requiring their uptake were all submitted later than May 2015.

For the six papers which submitted the appropriate CONSORT checklist, only one paper had complete consistency between the checklist and the published paper. The most concerning problems centred around items 8 and 9. Figure 5 summarises the inconsistencies found for each item. The evaluations for all papers can be accessed in Zenodo (21).

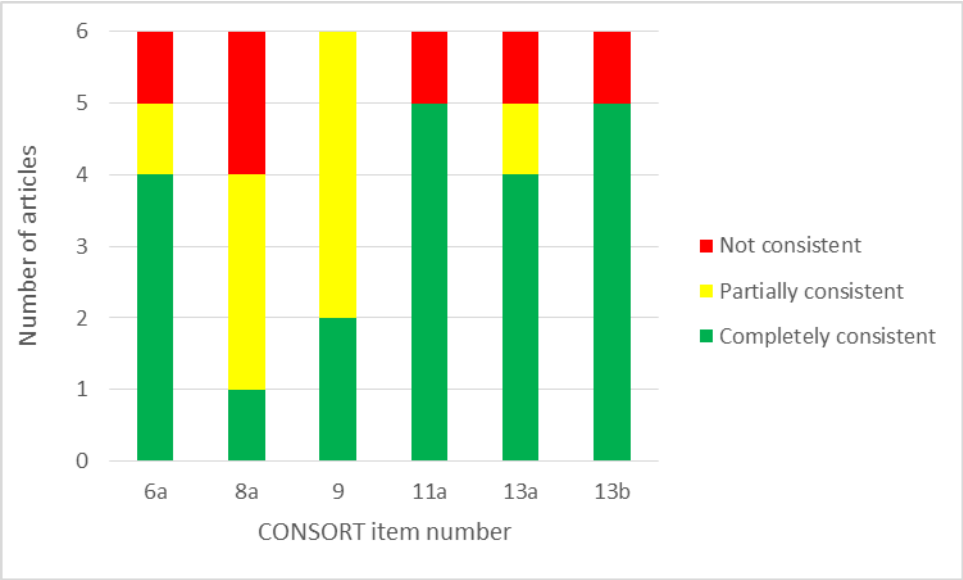


Figure 5: Reporting inconsistencies found for the six papers that used the appropriate CONSORT checklist. *Items:*

- *6a ("Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed"),*
- *8a ("Method used to generate the random allocation sequence"), 9 ("Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned"),*
- *11a ("If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how"),*
- *13a ("For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome"),*
- *13b ("For each group, losses and exclusions after randomisation, together with reasons"),*
- *17a ("For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)").*

For example, an inconsistency identified regarding CONSORT Item 9 ("*Mechanism used to implement the random allocation sequence, describing any steps taken to conceal the sequence until interventions were assigned*") was the following: authors claimed through the checklist that they had reported the item, however we found the paper cites "*The sequence of the test conditions was randomized for each participant by LB and KDK. A card was made for each possible sequence and a card was picked blindly for each participant*". This statement does not make clear how the authors implemented the random allocation sequence nor how they kept the assignment concealed. Picking a card does not guarantee that allocation used in the analysis has preceded treatment, neither allows readers to reproduce the mechanism used to implement the random allocation sequence. We show further illustrative examples of inconsistencies for the other CONSORT items in Table 5.

CONSORT item	Information reported	Evaluation	Evaluation rationale
6a: *Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed.*	*No information related to this item is visibly reported where referenced.*	Not consistent	Firstly, the outcomes are not clearly specified. Secondly, while the authors say in the abstract "Denture biofilm coverage was scored", this outcome is not mentioned in the location referenced in the checklist.
8a: *Method used to generate the random allocation sequence.*	"*Participants were randomly assigned to one of two parallel groups, in a 1:1 ratio*".	Not consistent	The method used to generate the random allocation sequence is not explicitly mentioned.
11a: *If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how.*	"*Patients were randomized by the study nurse, blinded from both the investigator and study participant*".	Not consistent	No information is provided about the degree of blinding of care providers and those assessing the outcomes.
13a: *For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome.*	"*Allocated to intervention (n=515) → Follow-up observations (n=925) - Study termination/Lost in follow up (n=364)*".	Not consistent	N is used for two different units: observations and participants. Moreover, authors report two different reasons simultaneously: study termination and lost to follow up.
13b: *For each group, losses and exclusions after randomisation, together with reasons.*	In the text: "*In the exercise condition, four participants dropped out during the intervention: reasons were injuries (n = 3) and migration (n = 1)*". In the flow diagram: "*Allocated*	Not consistent	The lost to follow up numbers reported are different in the text and the flow diagram.

	to exercise intervention (n=50) - Received allocated intervention (n=50) → Analysed: post-intervention (n=49) - follow-up (n=48)".		

Table 5: Examples of reporting inconsistencies

3.3. Why could reporting inconsistencies occur?

Among the numerous potential reasons for the presence of reporting inconsistencies, we underline two explanations. Firstly, it is possible that authors are not attentive to the requirements of CONSORT or, despite their efforts to be compliant with the requirements, they are struggling to interpret certain items or the level of detail that is required. Examples include:

- Item 8a ("*Method used to generate the random allocation sequence*"): this item was adequately reported in just one of the papers screened. Within this item there were various reasons for the inconsistencies, including: lack of thorough and complete reporting from the authors (see example for this item in Table 5) and the non-technical use of the term "random" ("The study nurse randomly opened a preformed envelope containing the allocated treatment regimen").
- Item 11a ("*If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how*"). The initial "If done" may have caused confusion about whether or not the authors have to report what groups of individuals involved in the trial were unblinded.

To avoid authors not disclosing the lack of blinding if the trial was unmasked, we suggest future versions of the CONSORT checklist to delete “if done”.

The misinterpretation of CONSORT is a major concern as it means that essential information regarding study conduct is miscommunicated. This is particularly relevant for Item 11a, as according to PRISMA Item 19 (152), when assessing the risk of bias of a study it is necessary to know whether patients, health care providers, data collectors, and outcome assessors are blinded or not.

Secondly, the issues described in this study might also lie with the reviewers and editors. It is possible that they are falsely reassured with regard to the reporting quality of the manuscripts, merely by the presence of a completed checklist. Moreover, the fact that the reporting inconsistencies persist throughout the editorial process might mean that editors and reviewers are not using RGs as a method to review manuscripts (125) although the CONSORT E&E document suggests that: “Readers, peer reviewers, and editors can also use CONSORT to help them critically appraise and interpret reports of RCTs”.

There are some considerations that may affect the generalisability of the results. Firstly, we only considered CONSORT among all existing RGs. We could expect similar reporting inconsistencies to also be frequent for other RGs that are less well-established than CONSORT. For example, it has been shown that requiring the submission a completed checklist of Animal Research: Reporting In Vivo Experiments (ARRIVE) is not enough to improve adherence in the context of a general scientific and medical journal (153). Secondly, we analysed a small sample of RCTs, which hinders the exact quantification of the reporting issues observed. Nevertheless, similar problems have been observed in a

larger sample of RCTs analysed in Chapter 5. Thirdly, we only considered three journals. As these are large medical journals that publish research across multiple specialties, we could expect these reporting issues to be present for all those medical journals that do not make sure that the submitted checklists are evaluated by the peer reviewers or the editorial staff.

3.4. Possible solutions

In an effort to take full advantage of requiring the submission of checklists, journals should consider clarifying their stance on whether the full checklists, or at least the core items of them, should be examined by editors or reviewers, or even by trained editorial assistants (127).

As the page numbers reported by authors in the checklist are not updated after the peer review process and the typesetting process, they do not correspond to the page where the information is placed in the published paper. Having to update the page numbers in the checklist from original submission to published paper could act as checkpoint for editors or reviewers to remind them to verify whether authors are appropriately reporting the key information in the latest version of the manuscript. An alternative solution could be to ask authors to address the section and the paragraph where the information corresponding to each item is reported. This would reduce the risk of overburdening the authors and could potentially help deter the misconception that these checklists are merely bureaucratic. Furthermore, making available the updated checklist could help systematic reviewers easily and quickly find the relevant information to assess the risk of bias of the studies included in the reviews (152).

Conclusions

Poor reporting critically affects the credibility and reproducibility of the methods and findings of RCTs. For these reasons, further exploration of methods that will obligate authors to consistently and accurately fulfil and submit CONSORT checklists is required. Moreover, journals should consider making clear whether the checklists should be examined by editors, peer reviewers, or by trained editorial assistants.

Chapter 4: A survey to explore the practicalities of the implementation of various editorial interventions to improve adherence to RGs

This chapter is based on the following published research paper:

- **Title**: A survey exploring biomedical editors' perceptions of editorial interventions to improve adherence to reporting guidelines (154)
- **Published in**: F1000, September 2019
- **DOI**: 10.12688/f1000research.20556.3
- **PubMed ID**: 31824668
- **Authors**: David Blanco, Darko Hren, Jamie J. Kirkham, Erik Cobo, Sara Schroter

4.1. Study objectives

This survey aimed to inform the future evaluation of interventions to improve adherence to RGs. In particular, we focused on interventions that can be implemented at various points in the editorial process.

Our specific objectives were to explore the perceived ease of implementation of different interventions and the potential effectiveness of these at improving adherence to RGs; to map the barriers and facilitators associated with these interventions; to determine possible solutions to overcome the barriers described, and to identify further editorial interventions that could be implemented and subsequently evaluated.

4.2. Methods

Participants

Purposive sampling was used to recruit biomedical editors that were expected to be knowledgeable and experienced in the topic we aimed to explore. We recruited participants not based on their representativeness of all medical journals but on the fact that they were "information-rich cases" (155).

Participants were sampled from three sources: (i) editors of journals that had published studies describing interventions to improve adherence to RGs identified in our scoping review (24), (ii) members of the MiRoR Network with current editorial positions and (iii) editors of the top-10 journals (based on impact factor) of BMJ Publishing Group which, apart from being one of the partner institutions of MiRoR, has published the main RGs (7,8,146,156) and has traditionally performed research to improve the transparency and quality of biomedical publications (157). The authors of this survey who met the eligibility criteria were excluded as potential participants.

Recruitment

The survey was only open to editors that we invited to participate. We contacted three editors (including the editors-in-chief) of each of the sampled journals, as well as individual editors from the group (ii) above. By replying to our invitation email, participants could suggest further editors that they considered could contribute to the survey. To contact editors not known to us we sought email addresses in the public domain. The survey was not advertised on any website.

Survey administration

The survey was administered by SurveyMonkey (158) and was open between 27 November 2018 and 24 February 2019. Participants were sent a personalised email inviting them to complete an online survey investigating their opinions about different editorial interventions to improve author adherence to RGs. Each invitation was tied to a unique email address. Two reminders to complete the survey were sent to non-responders at four and eight weeks after the initial mailing.

Participants could edit their responses while completing the survey. However, they could not re-enter the survey once it was completed as no two entries from the same IP address were allowed. We did not offer any incentives for completing the survey.

Response rates

We recorded the view rate of the invitation email (subjects opening the invitation email/subjects invited), the response rate (subjects completing the survey/subjects invited), and the completion rate (subjects completing the survey/subjects completing the first question of the survey).

Questionnaire development

Our previous scoping review (24) identified 31 interventions to improve adherence to RGs. For use in this survey, we chose a smaller subset of nine interventions that could be implemented during the editorial process as our focus was on journal editors' perceptions. These target different stakeholders:

A. Interventions targeting authors:

- A requirement for authors to submit a completed RG checklist with the specific page numbers where each item is addressed (**Intervention 1**)
- A requirement for authors to submit a populated RG checklist with text from their manuscript instead of page numbers (**Intervention 2**)
- A requirement for authors to highlight in the manuscript where each RG item is addressed (**Intervention 3**)
- A requirement for authors to include new subheadings within their manuscript corresponding to different RG items within the traditional IMRaD format (Introduction, Methods, Results, and Discussion) (**Intervention 4**)
- A requirement for authors on submission to use a freely available writing aid tool that guides authors through the RG checklist items, shows the key elements that need to be reported, and includes examples of adequate reporting (e.g. COBWEB) (**Intervention 5**)

B. Interventions targeting peer reviewers:

- Instruct peer reviewers to use the appropriate RGs when assessing a manuscript (**Intervention 6**)
- Instruct peer reviewers to scrutinise the completed RG checklist submitted by the authors and check its consistency with the information reported in the manuscript (**Intervention 7**)

C. Interventions targeting editorial staff:

    - An evaluation of the completeness of reporting by a trained editor (or editorial assistant), who would return incomplete manuscripts to authors before considering the manuscript for publication (**Intervention 8**)

D. Interventions targeting authors, peer reviewers, and editors:

    - Training for authors, peer reviewers, and editors on the importance, content, and use of RGs (e.g. The EQUATOR Network toolkits) (**Intervention 9**)

We pilot tested the draft survey questionnaire with two collaborators of the MiRoR project who currently hold editorial positions in biomedical journals. They were asked to review the survey for its clarity and completeness and to provide suggestions on how to improve its structure. Based on feedback from the pilot we decided not to include the intervention "Implementation of the automatic tool StatReviewer" (128) since participants were not aware of this software and stated that their perceptions would strongly depend on details about how it operates which are not publicly available.

The survey combined open and closed response questions to seek participants' perceptions of a series of interventions to improve authors' adherence to RGs that could potentially be implemented during the editorial process. The survey questionnaire (see Appendix B) is structured as follows:

- *Part 1: Current practice*. Participants were asked to describe the measures their journal currently takes to improve adherence to RGs.

- *Part 2: Perceptions of nine potential interventions*. Participants were asked to indicate on 5-point Likert scales (i) how easy it would be (or was) to implement these interventions at their journals (1-very difficult, 2-moderately difficult, 3-neither difficult nor easy, 4-moderately easy, 5-very easy) and (ii) how effective they thought the interventions would be (or was) at improving adherence to RGs if these were implemented at their journals (1-very ineffective, 2-moderately ineffective, 3-neither ineffective nor effective, 4-moderately effective, 5-very effective). We included images to clarify meanings and context to prompt participants to think about the benefits and drawbacks of the interventions. Free text boxes were included so participants could justify their responses.

- *Part 3: Identifying the barriers and facilitators.* Participants were asked to choose which intervention they considered potentially the most effective for their journal at improving adherence to RGs. They were asked to describe (i) why they thought that the intervention would be the most effective, (ii) what the main difficulties in implementing that intervention would be, and (iii) how they would try to overcome these difficulties.

- *Part 4: Further interventions.* Participants were asked for further suggestions of possible interventions, including modifications and combinations of the interventions previously discussed.

- *Part 5: Demographic questions*.

The survey was distributed over 18 pages with 1 to 3 items per page. We did not randomise the order of presentation of these items.

Data analysis

For quantitative data (Part 2 of the questionnaire), we used R version 3.6.0 (147). As these data were ordinal, we calculated medians together and the 1st and 3rd quartiles. We excluded from the analysis one questionnaire where the participant just opened the survey and left without answering any question. We did not exclude any questionnaire based on the amount of time that the participant needed to complete it.

For qualitative information, the lead investigator (DB) used the software program NVivo 12 (159). We mapped the barriers and facilitators for each of the interventions explored, as well as other key themes such as the incentives for the use of RG and the implementation of further editorial strategies. The initial mapping made by the lead investigator was discussed with another investigator (SS) and subsequently refined.

For Part 1 of the survey (*Current practice*) the unit of measure were the journals and therefore editors of the same journal were grouped. This was due to the fact that participants' answers represented an overarching policy and not an individual's opinion. For all other parts of the survey (Part 2 to Part 5), we analysed editors' responses independently, no matter what their journal was.

Ethics approval, informed consent and data protection

The Research Committee of the Governing Council of the Universitat Politècnica de Catalunya (UPC) granted ethical approval for this study (Reference EC 01, Date 2 May 2018).

In the invitation email, we informed survey participants that (i) the completion of the survey indicated consent to participate, (ii) they were free to stop and withdraw from the study at any time without providing a reason, (iii) the estimated time to complete the survey was 15 minutes, (iv) any identifiable information obtained in connection with this survey would remain confidential, and (v) the results would be submitted for publication and the anonymised dataset would be made publicly available in Zenodo (22). The original dataset was kept in a password-protected folder in Google Drive.

Reporting guidelines

We consulted the Checklist for Reporting of Results of Internet E-Surveys (CHERRIES) (160) and the Consolidated criteria for Reporting of Qualitative research (COREQ) (161) guidelines to produce this research report.

4.3. Results

Of the 99 editors invited, 42 opened the invitation (view rate 42%), and 24 completed the survey (response rate 24%) from the 25 who started it (completion rate 96%). The average time spent completing the survey was 15 minutes (SD = 8.5 minutes). Among the 24 participants who completed the survey, nine (37%) worked for seven different journals that had published studies on improving adherence to RGs, seven (29%) worked for five top-10 BMJ journals, four (17%) were members of the MiRoR Network that hold editorial positions in four journals, and a further four (17%) were suggested by other participants based on their expertise on the topic and were editors of three different journals. The 19 journals represented in the survey are listed in Table 6.

Journals that have published research on RGs	Trials
	The Lancet
	PLOS ONE
	BMC Medicine
	BMC Medical Research and Methodology
	Journal of Oral Rehabilitation
	Journal of Clinical Epidemiology
Journals that belong to BMJ top-10	The BMJ
	Archives of Disease in Childhood
	BMJ Open
	BMJ Quality & Safety
	British Journal of Sports Medicine
MiRoR Network members' journals	Clinical Chemistry
	Systematic Reviews
	Research Integrity and Peer review
	Journal of Global Health
Other journals	F1000
	BMJ Open Science
	Scientific Reports

Table 6: Journals represented in the survey

Participants had a variety of editorial roles (editor-in-chief, senior editor, associate editor or others). Most of them were involved in manuscript decision-making and had less than 15 years of experience as journal editors (see Table 7). The anonymised responses from all 24 participants can be accessed in Zenodo (22).

N=24		
Current position	Working full time as a journal editor	8 (33%)
	Working part time (equal or more than 0.5 of their time) as a journal editor	1 (4%)
	Working part time (less than 0.5 of their time) as a journal editor	14 (59%)
	Other (Volunteer editor)	1 (4%)
Editorial role	Editor-in-chief	10 (41%)
	Senior editor	4 (17%)
	Associate editor	4 (17%)
	Other (Editorial director, Technical editor, Assistant editor)	6 (25%)
Involvement in manuscript decision-making	Yes	22 (92%)
	No	2 (8%)
Years of experience as a journal editor	<5	8 (33%)
	5-15	12 (50%)
	15-25	3 (13%)
	>25	1 (4%)

Table 7: Demographic characteristics of the 24 participants

Current practice

Respondents worked at 19 journals. Most respondents' journals (11/19, 58%) request authors to submit a completed RG checklist with page numbers indicating where the items are addressed when they submit their manuscript. A further seven (37%) instruct but do not request authors to do it, and one (5%) does not request or instruct authors.

Among the journals requesting the submission of checklists, four (4/11, 36%) also explicitly ask peer reviewers to use the completed RGs when assessing manuscripts, one (1/11, 9%) asks peer reviewers general questions about the completeness of reporting, and one performs an evaluation of the completeness of reporting by a trained editor using RGs before the initial decision is made on the manuscript. We observed no incongruences between the answers of editors from the same journal. Some respondents mentioned that in their journals (n=4) the interventions described were only applicable to the study types corresponding to the most established RGs (CONSORT, PRISMA (8), STROBE (7)) for trials, observational studies and systematic reviews respectively.

Perceptions of nine potential interventions

The mean scores for perceived ease of implementation and potential effectiveness for each intervention are shown in Figure 6.

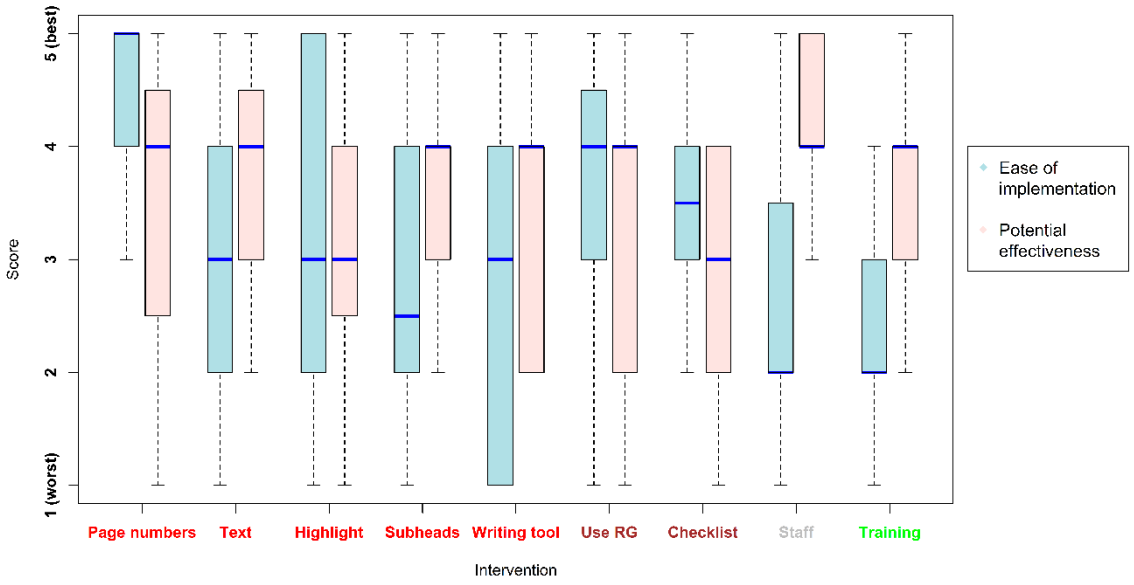


Figure 6: Scores for perceived ease of implementation and potential effectiveness*. Box plots show the 1st, 2nd (medians, represented as blue horizontal lines), and 3rd quartiles of the data. The whiskers extend up to 1.5 times the interquartile range from the top (bottom) of the box to the furthest datum within that distance. Interventions whose names are shown in red target authors, those in brown target peer reviewers, the one in grey target editors or administrative staff and the one in green targets all.*

The two most common interventions were considered the easiest ones to implement: the median scores (1^{st}, 3^{rd} quartiles) for requesting authors to submit checklists with page numbers (Intervention 1) and for asking peer reviewers to use RGs (Intervention 6) were 5 (Q_1: 4, Q_3: 5) and 4 (Q_1: 3, Q_3: 5), respectively. By contrast, interventions related to training (Intervention 9), editor involvement in checking completeness of reporting (Intervention 8) and reformatting of the text based on RG requirements (Intervention 4, Intervention 5) were considered the most difficult to implement.

An evaluation of the completeness of reporting by a trained editor was considered the most effective intervention at improving adherence to RGs (Median: 4, Q_1: 4, Q_3: 5) and

the two targeting peer reviewers (Interventions 6 and 7) were perceived as being the least effective (Median: 4, Q_1: 2, Q_3: 4; Median: 3, Q_1: 2, Q_3: 4).

Identifying the barriers and facilitators

This section presents the perceived barriers and facilitators of the interventions considered and editors' suggestions for making the interventions more effective. Appendix C shows a full description of these.

A) Interventions targeting authors (Interventions 1-5)

The main barriers associated with all of the interventions targeting authors was that authors have to state their adherence to the relevant RG and this does not equate to actual compliance. Moreover, it is resource intensive for journals to check that these requirements are appropriately met by authors. Some editors highlighted that Interventions 3, 4, and 5 would involve special formatting of the submitted manuscript, which could be cumbersome for authors given that manuscripts are often submitted to multiple journals with different formats before being accepted. This is particularly relevant for journals with high rejection rates as it could cause frustration for authors. Some participants mentioned logistical issues as their journal's manuscript tracking system is not set up to accommodate these interventions. In addition, changes in the manuscript's format could be incompatible with the journal's house style.

Intervention 1 was generally considered quick and straightforward for authors, but three participants indicated that there is published empirical evidence of little effectiveness if the checklist is not assessed by a trained editor or administrator (14,16,144,153).

As Interventions 3, 4, and 5 force authors to tailor the manuscript to RG requirements, participants reported that these could make editors' and peer reviewers' jobs easier as the manuscript would be better structured. Importantly, readers would also be able to locate information more easily. Some editors pointed out that, to make these interventions effective, journals would need to provide templates to authors or to integrate these interventions in the submission system. However, some of these interventions (Interventions 2 and 5) were seen as more effective if they were implemented earlier on in the research process, prior to writing the manuscript.

B) Interventions targeting peer reviewers (Interventions 6, 7)

Most respondents were negative about the potential effectiveness of implementing the two interventions targeting peer reviewers (Intervention 6 and 7) as they felt these would create too much additional work for reviewers. Participants were concerned that the quality of peer review could be compromised as reviewers are not expected to focus on reporting issues but on providing an expert view on the importance, novelty and relevance of the manuscript. Furthermore, peer reviewers may not know which RGs to use and, even if they do, the effectiveness would be dependent on their willingness to use RGs and their expertise in applying them. Several participants indicated that this work should be delegated to paid editorial staff.

C) Interventions targeting editorial staff (Intervention 8)

This intervention was considered difficult to implement but potentially effective. The main facilitating factor for its successful implementation was that it is performed by a paid or trained professional, which lends credibility to the intervention, reduces the

workload of unpaid peer reviewers, and avoids authors overclaiming adherence. The main barriers outlined for this intervention were (i) the budget issues the journal would need to face to train or hire additional editorial staff that could perform the evaluation, especially if the journal receives a large volume of manuscripts, (ii) the editorial delays it may cause, and the (iii) the potential inefficiency of assistant editors or administrators having to delegate decisions in case of doubt, given that sometimes assessing completeness of reporting is a subjective task.

To make this intervention more feasible for journals, editors suggested that the completeness of reporting evaluation could be performed only for manuscripts that are sent out for peer review and, it could be focused on a few core items (different for each RG) that would enable reproducibility. If this intervention was implemented in a journal that requires the submission of a completed checklist, editors could take advantage of the checklist to locate information.

D) Interventions targeting authors, peer reviewers and editors (Intervention 9)

Training was seen as a potentially effective intervention but difficult to implement. Some participants highlighted that training with follow up sessions would be resource intensive for journals, and especially difficult to enforce. One participant mentioned that credits, such as the Continuing Medical Education (CME) credits (162), could be used to recognise hours of training. The fact that sometimes the editorial staff is based in different places and zones makes it crucial to consider flexible forms of training, such as online courses. As an example, the EQUATOR Network Toolkits section provides resources for authors, peer reviewers and journal editors (163). However, some

participants emphasised that training should also be delivered by research institutions and medical centres.

Further interventions and incentives for authors and journals

When asked about further potentially effective interventions that were not discussed in the survey, some editors mentioned StatReviewer, a reading tool that automatically assesses adherence to RGs and is currently under evaluation (128). Other respondents also mentioned the possibility of combining some of the interventions discussed in the survey, such as requiring the submission of checklists and trained editors assessing the responses with the information reported in the manuscript.

Moreover, several incentives for authors were listed, including (i) discounts on article processing charges (APCs) for authors that comply with RG requirements, (ii) academic institutions including RG use in the promotion and tenure files, and (iii) credits (such as CME credits (162)) to recognise hours of training on the use of RGs. Journals could also be encouraged to implement certain interventions if (i) there is empirical evidence that these interventions actually improve the reporting quality of the papers or (ii) publishers or the International Committee of Medical Journal Editors (ICMJE) mandate these as a condition of submission to their journals. Even if some of these interventions are proven to be effective, some respondents reported that it is essential to convince publishers that improving the quality of reporting is a worthy investment to resource.

4.4. Discussion

We surveyed biomedical journal editors with experience and interest in the topic of improving authors' adherence to RGs. We aimed to explore their perceptions of practical aspects of the implementation of different interventions to improve adherence to RGs.

Several messages arise from this study. First of all, most editors agreed that the most effective way to improve adherence to RGs is for journals to involve trained editors or administrative staff. Interventions targeting these stakeholders were considered to be difficult to implement for most journals, either because of logistic or resource issues. However, improving the performance of editorial staff is critical (142) and has been shown to have a positive impact on completeness of reporting in the context of a dentistry journal (126). To make these types of interventions more feasible, journals could implement them only for manuscripts that are sent out for peer review. The editorial staff could also take advantage of the RG checklists submitted by authors, that could be automatically populated with text using specific software such as the tool proposed by Hawwash et al. (164).

Most editors considered that checking reporting issues is beyond the role of peer reviewers. Given the voluntary nature of peer review, requiring reviewers to use RGs would cause an additional workload that could compromise the overall quality of the reviews. If checking reporting issues becomes a standard exercise for peer reviewers, some editors are concerned that peer reviewers may be less likely to comment on important aspects of a manuscript, such as its novelty, clinical interest or implications. Furthermore, as finding peer reviewers is becoming increasingly difficult for editors (165), these requirements could make them even less willing to review papers.

Additionally, some editors considered that the average peer reviewer does not have enough expertise to go over RG requirements.

We observed that the interventions perceived as potentially most effective at improving adherence to RGs appear to be more difficult to implement. Conversely, the most common strategies seem to have been implemented based on their feasibility and not on their potential to improve completeness of reporting. This could be one of the reasons why they have failed to achieve the desired results (14,16,144,153). Some of our respondents insisted that a key element is that journals, universities, and medical institutions find ways to incentivise author's compliance with RGs. At the same time, the scientific community needs to find ways to convince publishers that improving the quality of reporting is a worthy investment so that publishers can encourage their journals to adopt strategies to boost completeness of reporting. A recent article indicates that implementing RGs through the editorial process may increase the number of citations to the research reported (166).

A common observation by the survey participants was that the effectiveness of the interventions proposed could depend on the types of articles considered. While RGs for RCT protocols, RCTs or systematic reviews are more established, some others, including most RG extensions, are not well known to the stakeholders involved in the publication process. For this reason, it is important for journals to be clear in their "Instructions for Authors" on what RGs they mandate.

It is noteworthy to mention that, regardless of how checklists are implemented in the editorial process and who has to engage to make the interventions successful, the evaluation of completeness of reporting is a subjective task. This is mainly due to the

fact that RGs are not originally designed as evaluation tools but as guidance for authors on how to report their research. For this reason, evaluators could sometimes have different views on whether authors are providing enough information to consider that certain RG items are adequately reported.

This study is subject to several limitations. The response rate was low (24%). Researchers in health science have witnessed a gradual decrease in survey participation over time (167), especially among health professionals due to the demanding work schedules and increasing frequency of being approached for surveys (168). Some recent surveys in the field of peer review show even lower response rates (10-20%) among researchers, peer reviewers and readers (169,170). It is also noteworthy that we took a pragmatic approach to identify relevant editors and the sample was small due to not many having conducted or published research on improving adherence to RGs. Whilst n=24 is a small number, the detailed and rich qualitative responses that we received showed a high level of engagement with the topic. Despite having the option to increase the sample size by contacting more editors at a lower level of hierarchy in the journals we targeted, we decided not to do it based on the response rate of the survey. That approach would have changed our sampling frame and we would potentially have had less experienced editors commenting. We took that decision as the purpose of the survey was to tap into the experience of those who had tried interventions or had shown interest in this area, instead of seeking a representative sample of editors.

Connected with this, we could expect survey participants to be more prone to adopt interventions than general biomedical editors. However, their experience could also make them more critical of certain strategies that appear to be more effective than they

actually are. This could be the case for the intervention of requesting authors to submit checklists on manuscript submission, which has become popular among medical journals despite having little or no impact on completeness of reporting (14,16,144,153). Editors with less experience of editorial strategies to improve adherence to RGs might expect authors and peer reviewers to respond to certain interventions in a different way than they would do.

We encourage researchers to perform further evaluations of interventions in collaboration with biomedical journals, such as the RCT reported in Chapter 5. Our study aims to evaluate the effect on completeness of reporting of a trained researcher assessing during peer review the consistency between the CONSORT checklists submitted by authors and the information reported in the manuscript, and providing authors with a report indicating any inconsistencies found.

Providing high quality evidence of the effectiveness of different interventions at improving adherence to RGs and discussing how to make them less burdensome are key aspects needed to convince all stakeholders that this effort is worth it.

Conclusions

Biomedical journal editors with experience and interest in the topic of improving authors' adherence to RGs generally believed that engaging trained professionals in the process of checking adherence would be the most effective, yet moderately resource intensive, editorial intervention. Also, they thought that standard peer reviewers should not be asked to check RG requirements.

Future evaluations of interventions to improve adherence to RGs can take into account the barriers, facilitators, and incentives for implementing editorial interventions that are described in this survey.

Chapter 5: A randomised controlled trial to evaluate the impact of an editorial intervention to improve the adherence to CONSORT

This chapter is based on the following research paper:

- **Title**: Effect of an editorial intervention to improve the completeness of reporting of randomised trials: a randomised controlled trial (171)
- **Published in**: BMJ Open, May 2020
- **DOI**: 10.1136/bmjopen-2020-036799
- **Authors**: David Blanco, Sara Schroter, Adrian Aldcroft, David Moher, Isabelle Boutron, Jamie J Kirkham, and Erik Cobo

5.1. Study objectives

This RCT assessed whether an editorial intervention performed by a researcher with expertise in CONSORT improved the completeness of reporting of the trials submitted to BMJ Open, a general medical journal. The intervention consisted of an evaluation of completeness of reporting of eight core CONSORT items using the submitted checklist to locate information, and the production of a report containing specific requests for authors based on the reporting issues found, provided alongside the peer review reports. This experiment was carried out as part of BMJ Open's quality improvement programme.

5.2. Methods

Trial design and study setting

This was a two-arm parallel randomised trial (1:1 allocation ratio) conducted in collaboration with BMJ Open, an open-access general medical journal (published by the BMJ Publishing Group) that requests the submission of completed CONSORT checklists for RCTs. Prior to recruitment, we registered the study in ClinicalTrials.gov with the identifier NCT03751878 and uploaded the study protocol (172).

Eligibility criteria

Manuscripts were eligible for inclusion if (i) they were original research articles reporting the results of an RCT submitted to BMJ Open, (ii) they had passed the first editorial filter and had been subsequently sent out for peer review, and (iii) authors of these manuscripts had provided a completed CONSORT checklist as part of the submission process. Apart from the standard two-arm parallel RCTs, which are covered by the standard CONSORT guidelines (149), we also included RCTs that require the use of the official CONSORT extensions for different design aspects (cluster (11), non-inferiority and equivalence (173), pragmatic (151), N-of-1 trials (174), Pilot and feasibility (10), and within person trials (175)) and intervention types (herbal (176), non-pharmacologic (150), acupuncture (177) and Chinese herbal medicine formulas (178)) in all areas of clinical research. We excluded (i) studies that claimed to be RCTs but used deterministic allocation methods, and (ii) secondary trial analysis studies.

Interventions

We designed a three-step intervention based on the results of our previous work (24,154) ensuring no disruption to usual editorial procedures. Firstly, DB assessed completeness of reporting of eight core CONSORT items (see the following paragraph) using the submitted checklist to locate the information corresponding to each item. Secondly, DB produced a standardised report containing precise requests to be addressed by authors. This report included a point by point description of the reporting issues found, requests to the authors to include the missing information (see example in Box 2), as well as examples extracted from the CONSORT E&E document (149). Finally, DB uploaded the report to the manuscript tracking system of the journal (ScholarOne) to make it accessible to the manuscript handling editor, who included this additional report in the decision letter to authors alongside the standard peer review reports. Manuscripts randomised to the control group underwent the usual peer review process. In Figure 7, we display a schema of the study design.

The intervention was focused on eight core CONSORT items (see Box 3) which are essential for researchers evaluating the risk of bias of RCTs when conducting systematic reviews (148) and which are usually poorly reported (127).

Please, make the following revisions:

- For CONSORT Item 8a ("*Method used to generate the random allocation sequence*"), please report the exact method you used to generate the random allocation sequence.
    - Example from CONSORT: "*Randomization sequence was created using Stata M.N (StataCorp, College Station, TX) statistical software*".
- For CONSORT Item 11a ("*If done, who was blinded after assignment to interventions and how*"), please specify in "Trial design and setting" who was blinded in the study and do not just state that it was a double-blind randomised trial.
    - Example from CONSORT: "*Whereas patients and physicians allocated to the intervention group were aware of the allocated arm, outcome assessors and data analysts were kept blinded to the allocation*".

Box 2: Report reflecting the reporting inconsistencies

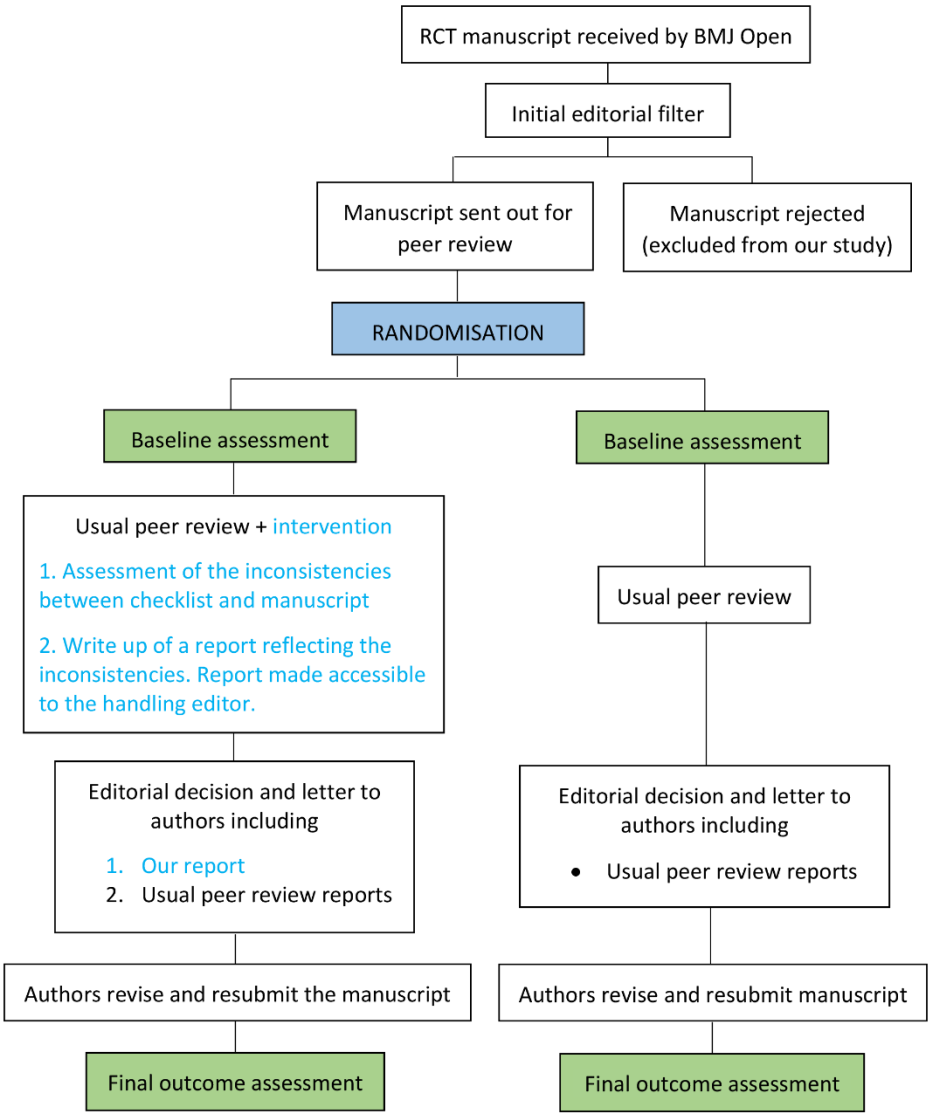


Figure 7: Schema of the study design

Five items in the methods section:

- **Item 6a** ("*Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed*")
- **Item 8a** ("*Method used to generate the random allocation sequence*")
- **Item 9** ("*Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned*")
- **Item 11a** ("*If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how*")
- **Item 11b** ("*If relevant, description of the similarity of interventions*")

Three items in the results section:

- **Item 13a** ("*For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome*")
- **Item 13b** ("*For each group, losses and exclusions after randomisation, together with reasons*")
- **Item 17a** ("*For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)*")

Box 3: Core CONSORT items considered.

We considered an item as adequately reported if all subparts of it were adequately reported, according to the CONSORT E&E document (149) and the corresponding E&E documents for the extensions considered. For example, for CONSORT item 6a ("*Completely defined pre-specified primary and secondary outcome measures, including*

how and when they were assessed"), we required the following subparts to be adequately reported: A) identified and completely defined primary and secondary outcomes, B) analysis metric and methods of aggregation for each outcome, and C) time points for each outcome.

The items corresponding to CONSORT extensions were assessed in addition to the standard CONSORT items. For example, we expected authors of a cluster randomised trial evaluating a pharmacologic treatment to be using the standard CONSORT checklist for all eight items and the cluster extension for items 6a, 9, 13a, 13b, and 17a. In contrast, the items requested by the Pilot and Feasibility extension substituted the standard CONSORT items, as specified in its E&E document (10). Once the recruitment had begun, we decided to discard the extension for non-pharmacologic interventions as it was not being requested by the editors, nor sent by authors.

For items reported as N/A in the CONSORT checklist, we considered them as:

- Adequately reported if (i) the item did not apply and therefore it did not have to be reported, and (ii) the item applied and it was actually reported although the page number was not given.
- Inadequately reported if the item did apply but it was not adequately reported.

We also applied some rules on how to deal with certain aspects of specific items:

- **Item 8a**: inadequately reported if authors have reported this information elsewhere but not in the main body of the article. According to CONSORT, "*it is important that information on the process of randomisation is included in the*

body of the main article and not as a separate supplementary file; where it can be missed by the reader".

- **Item 11a**: adequately reported if blinding was not performed and authors explicitly said so, and inadequately reported if blinding was assumed to be not performed and authors did not mention it in the manuscript.
- **Item 13a** and **item 13b**: the corresponding information could be included either in the text or in the flow diagram. If information was only included in the discussion, it was considered as inadequately reported.
- **Item 17a**: adequately reported if there was a correspondence between the outcomes in the results section and the ones listed in the methods section (and therefore evaluated in Item 6a).
- **Extension of Item 17a** for Pilot and Feasibility trials ("*For each objective, results including expressions of uncertainty (such as 95% confidence interval) for any estimates. If relevant, these results should be by randomised group*"): we did not expect authors to report the effect sizes but the results (plus expressions of uncertainty) for each objective.

Outcomes

- **Primary outcome**: Mean score for completeness of reporting, defined as the mean number of adequately reported items in the first revised manuscript (0 to 8 scale).
- **Secondary outcome**: Proportion of manuscripts where each item was adequately reported.

For each of the manuscripts in the intervention group, we recorded the amount of time that took the lead investigator to perform the assessment of reporting inconsistencies and to produce the report.

In the design phase of the study, we considered two potential scenarios where included manuscripts could potentially be lost to follow-up: (i) when editors rejected a manuscript after peer review, and (ii) when authors did not return the revised manuscript within the period requested by the handling editor after a "Minor revision" or "Major revision" editorial decision (14 and 28 days, respectively, plus, if necessary, the extra time that the editor considered appropriate). In the "Statistical methods" section, we report the methods used to impute the study outcomes for lost to follow-up articles.

Outcome evaluation was performed independently and in duplicate by two senior researchers (EC, JJK) who were blinded to manuscript allocation and had experience as authors and reviewers of RCTs. They also assessed outcomes at baseline. In cases where a manuscript was rejected after the first round of peer review, assessors could only evaluate it at baseline. However, they were not aware of the fate of that manuscript until after they had completed that evaluation.

We performed the outcome assessment process as follows: first, DB divided the 24 included manuscripts into 4 batches of 6 manuscripts. Every time DB detected in the submissions report (see "Pilot work" section) that all 6 manuscripts of each batch had been revised by authors, DB first made available to the outcome assessors the submitted version of the manuscript (version 1). Assessors had to complete the evaluation form for each manuscript independently and in duplicate. This form included the CONSORT

extensions to be used. Assessors could explicitly indicate in it that they wanted to discuss a specific item with the other assessor. Once they were done with all manuscripts' version 1, DB informed them of the discrepancies between their evaluations, which were resolved by consensus. Afterwards, DB shared the manuscript revised by the authors (version 2) and we repeated the outcome evaluation process. This process was done for the 4 batches of 6 manuscripts.

For each of the manuscripts in the intervention group, we also recorded the amount of time it took the lead investigator to perform the intervention.

Harms

We analysed whether our intervention caused the following unintended effects: higher proportion of manuscript rejections after the first round of peer review and delays in the submission of the revised manuscripts by authors.

Pilot work

To inform the sample size calculation, the lead investigator assessed 12 randomly selected RCTs published in BMJ Open between April 2018 and September 2018. The proportions of adequately reported items observed in these manuscripts were used to estimate the scores for completeness of reporting of the manuscripts in the control group (usual peer review).

Furthermore, outcome assessors (EC and JJK) practised the evaluation of completeness of reporting by assessing six of the 12 RCTs mentioned above.

Power analysis

According to the assessment described in the "Pilot work" section, the estimated probabilities that manuscripts in the control group adequately reported 0, 1, 2,…, and 8 items were 0, 0, 0, 0, 0, 0.17, 0.33, 0.33, and 0.17, respectively. With the intervention, we aimed to bring this distribution to 0, 0, 0, 0, 0, 0, 0, 0.5, 0.5. In other words, manuscripts in the intervention group were expected to be adequately reporting 7 or 8 items 50% of the time, respectively.

In order to relax the strong required assumptions behind using a t-test for a reduced sample size, we used bootstrapping, a simple yet powerful non-parametric technique (33). First, given the probability distributions mentioned above, we performed 10.000 simulations of the scores of n manuscripts. We resampled each of these simulations 10.000 times in order to calculate the 95% CI of the mean difference between groups. Finally, we calculated the study power by counting for how many of the 10.000 simulations the lower limit of this 95% CI was over 0.

Choosing a sample size of n = 24 manuscripts (12 per arm) and following the procedure above gave us 90% power (alpha = 0.05, two-tailed). The R code used for the power analysis, as well as the other R codes used in this study, can be found in Zenodo (23).

Randomisation and blinding

Prior to recruitment of manuscripts, DB screened automated reports listing original research submissions to BMJ Open on ScholarOne, daily, including their identification (ID), date of submission, title, abstract, and different parameters related to their peer review status. RCTs were identified for possible inclusion based on the title and abstract

and then checked against our eligibility criteria until the desired sample size was achieved.

Every time a manuscript met our eligibility criteria, DB introduced its ID into an R Shiny application (179) created by a senior statistician (JAG) (23), which randomised the manuscript to the intervention or the control group (1:1 allocation ratio, blocks of 4). Manuscripts were stratified according to whether there was an applicable CONSORT extension for that study or not. To avoid allocation bias, each ID could only be introduced once.

As part of the usual submission process, all authors are informed that BMJ Publishing Group has a quality improvement programme and their manuscript might be entered into a study. However, authors of included manuscripts were not explicitly informed that their manuscripts were part of an RCT.

Outcome assessors were blinded to allocation and to each other's evaluation. Handling editors of the included manuscripts and the investigator performing the intervention (DB) were not blinded.

Statistical methods

We carried out statistical analysis using R version 3.6.0 (147).

For the primary outcome, we adjusted a linear regression model with the baseline score of the manuscript as the only covariate. We calculated the 95% confidence interval using bootstrapping (23).

The main analysis of the primary outcome was intention-to-treat: all manuscripts were included in this analysis regardless of whether they were lost to follow-up. We imputed the scores of lost to follow-up manuscripts with a value of 8-b, where b was the baseline score of the manuscript. This imputation strategy aimed to reflect the fact that rejecting RCTs of low baseline quality could be considered an editorial success. In addition, we assessed the sensitivity of the results by carrying out a complete case analysis and analysing the best case (manuscripts in the intervention group reached the maximum score and controls did not improve) and worst case (manuscripts in the intervention group did not improve and controls reached the maximum score) scenarios.

We did not plan any subgroup analysis (see protocol (172)) and so none are reported.

Amendments to the protocol

The last criterion listed above (iii) authors of the manuscripts had provided a completed CONSORT checklist) was not included in the first version of the protocol but we implemented it before recruitment started. The reason was that, despite that the submission of the CONSORT checklist for trials is mandatory, we observed that handling editors were occasionally overlooking this requirement and sending out manuscripts of trials for peer review that did not include one. Secondly, we initially used a t-test to calculate the study power and planned to use it for the primary outcome analysis. However, for the reasons described in the "Power analysis" section we used a bootstrap approach and the study power increased from the 85% stated in the protocol to 90%. Thirdly, masked to study results, we decided to adjust for the baseline scores in the primary outcome analysis. Given the reduced sample size of this RCT, we aimed to avoid that the differences between groups were random and due to the potential imbalance

in baseline score between groups. Finally, we added a best- and worst-case scenario analysis to assess the sensitivity of the primary outcome results.

Reporting guidelines

We report this manuscript in accordance to CONSORT 2010 (146).

5.3. Results

Between 31 October 2018 and 4 April 2019, we screened 62 manuscripts that described RCTs submitted to BMJ Open. Among these, we excluded 38 either because they were rejected without peer review (n = 34) or because the authors did not provide the CONSORT checklist (n = 4). We randomised the remaining 24 to the intervention (n = 12) or control (n = 12) groups. Six manuscripts (25%) were lost to follow-up (intervention n = 3, control n = 3) as they were rejected after the first round of peer review and therefore not returned to authors for revision (scenario (i) in "Outcomes" section). No manuscripts were lost to follow-up in scenario (ii) as all authors returned the revised manuscripts within the given time. Therefore, 18 manuscripts (intervention n = 9, control n = 9) were revised by authors. Figure 8 shows the flow diagram of the study.

Most manuscripts (n = 19, 79%) required at least one extension: non-pharmacologic (intervention n = 10; control n = 8), pilot and feasibility (n = 3; n = 4), cluster (n = 2; n = 1). Table 8 displays the baseline characteristics of the included manuscripts.

The mean (SD) baseline score for completeness of reporting (0 to 8 scale) prior to peer review in the intervention (n = 12) and control (n = 12) groups was 4.35 (1.88) and 4.85 (1.79), respectively. The mean (SD) baseline score of the manuscripts that later passed

the first round of peer review (n = 18) were much more complete (scores almost double) than those that were rejected after the first round of peer review (n = 6): 5.23 (1.35) versus 2.68 (1.75).

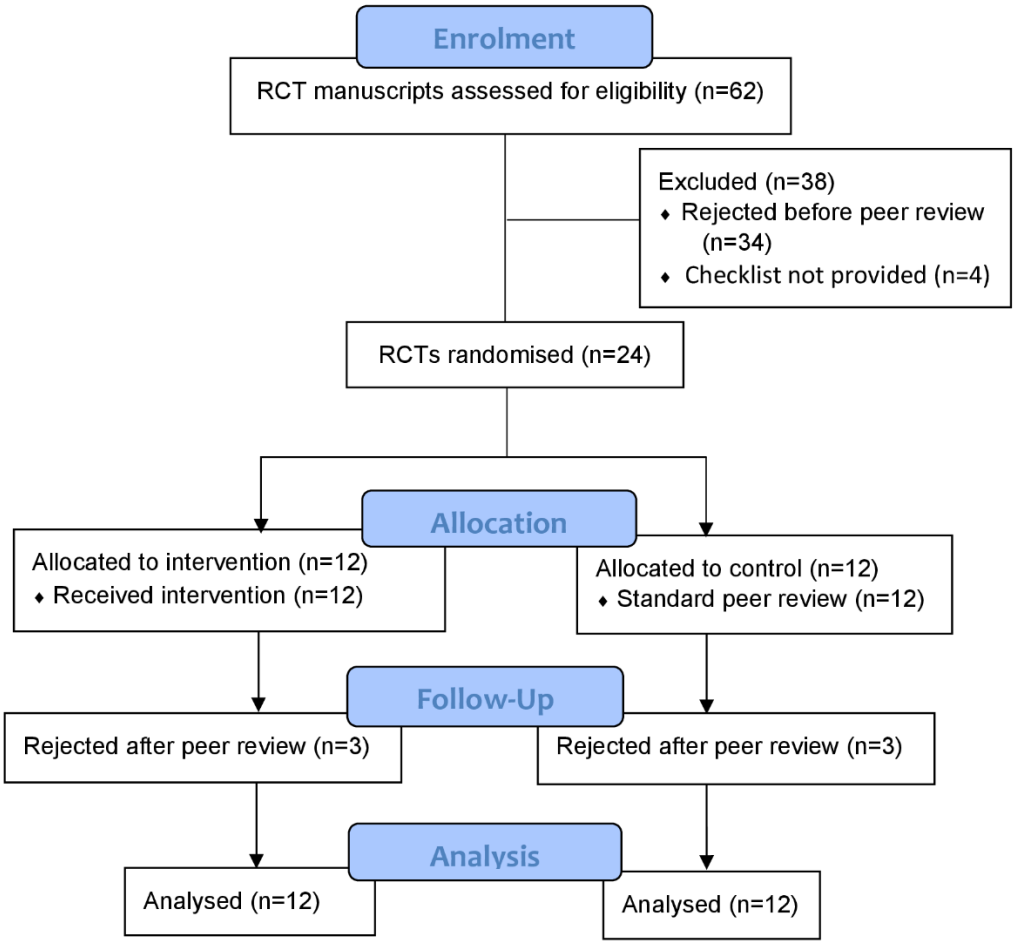


Figure 8: CONSORT flow diagram

			Intervention (n=12)	**Control (n=12)**
Study Design	Standard parallel-group		7 (58%)	7 (58%)
	Cluster		2 (17%)	1 (8%)
	Pilot & feasibility		3 (25%)	4 (33%)
Type of intervention	Pharmacologic		2 (17%)	4 (33%)
	Non-pharmacologic		10 (83%)	8 (67%)
		Behavioural	4 (33%)	3 (25%)
		E-health & tele-health strategies	3 (25%)	2 (17%)
		Medical devices	2 (17%)	1 (8%)
		Surgery	0 (0%)	1 (8%)
		Others	1 (8%)	1 (8%)
Single- or multi-centre	Single-centre		8 (67%)	5 (42%)
	Multi-centre		4 (33%)	7 (58%)
Number of participants	≤ 50		5 (42%)	2 (17%)
	> 50 & ≤ 100		3 (25%)	7 (58%)
	> 100		4 (33%)	3 (25%)
Registered in a trial registry	Yes		11 (92%)	11 (92%)
	No		1 (8%)	1 (8%)

First author's affiliation	Asia	3 (25%)	3 (25%)
	UK	3 (25%)	5 (42%)
	Europe	2 (17%)	3 (25%)
	USA	2 (17%)	0 (0%)
	Australia	2 (17%)	0 (0%)
	Brazil	0 (0%)	1 (8%)
Sponsorship	Investigator-initiated	12 (100%)	10 (83%)
	Industry-initiated	0 (0%)	2 (17%)

Table 8: Baseline characteristics of the included RCTs.

Primary outcome

For the intention-to-treat analysis (n = 24), the manuscripts that received the intervention were more completely reported than the ones that underwent the standard review process: intervention group (mean: 7.01; SD: 1.47) versus control group (mean: 5.68; SD: 1.43). After adjusting for the baseline score, the mean difference in scores between the two groups was 1.43 (95% CI: 0.31 to 2.58); manuscripts in the intervention group reported on average 1.43 (out of 8) items more adequately than those receiving the standard peer review. Regarding the sensitivity analysis, for the complete case (n = 18) the mean (SD) scores for the intervention and control groups were 7.45 (1.00) and 5.90 (1.35), giving an adjusted difference of 1.75 (95% CI: 0.80 to 2.75). The best- and worst-case scenario analysis (n=24) lead to adjusted differences of

2.62 (95% CI: 1.49 to 3.65) and 0.03 (95% CI: -1.45 to 1.63) respectively. Table 9 summarises these results.

Outcome	Intervention group Mean (SD)		Control group Mean (SD)		Mean difference in final scores* (95% CI)
	Baseline	**Final**	**Baseline**	**Final**	
Completeness of reporting (0 to 8 scale) with imputation (n = 24)	4.35 (1.88)	7.01 (1.47)	4.85 (1.79)	5.68 (1.43)	1.43 (0.31 to 2.58)
Completeness of reporting (0 to 8 scale) without imputation (complete case analysis, n = 18)	5.01 (1.32)	7.45 (1.00)	5.46 (1.41)	5.90 (1.35)	1.75 (0.80 to 2.75)
Completeness of reporting (0 to 8 scale) in the best-case scenario (n=24)	4.35 (1.88)	7.59 (0.89)	4.85 (1.79)	5.18 (1.89)	2.62 (1.49 to 3.65)
Completeness of reporting (0 to 8 scale) in the worst-case scenario (n=24)		6.18 (2.61)		6.43 (1.49)	0.03 (-1.45 to 1.63)

*Adjusted for baseline score.

Table 9: Primary outcome results.

Figure 9 shows the evolution of the 18 manuscripts that were revised and resubmitted. From the nine manuscripts in the intervention group, six of them achieved the maximum score and another two improved. In contrast, the only manuscript in the control group that reached the maximum score already had that score at baseline. Three manuscripts in the control group slightly improved (1, 1, and 2 points respectively). We identified that 3 out of 4 of these improvements were the result of comments made by the standard peer reviewers, rather than the authors themselves.

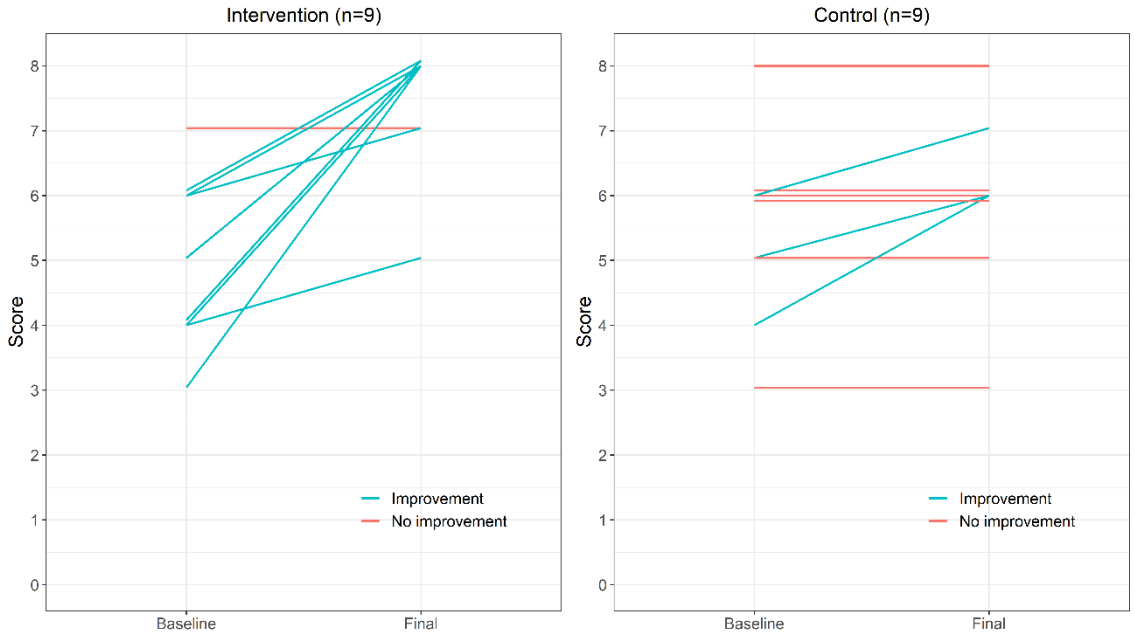


Figure 9: Evolution of the scores scores for all manuscripts that passed the first round of peer review (n=18).

Secondary outcome

Figure 10 displays the proportions of manuscripts where each CONSORT item was adequately reported. We observed the main differences favouring the intervention group in items 6a (Outcomes), 9 (Allocation concealment mechanism), 11a (Blinding), and 17a (Outcomes and estimation).

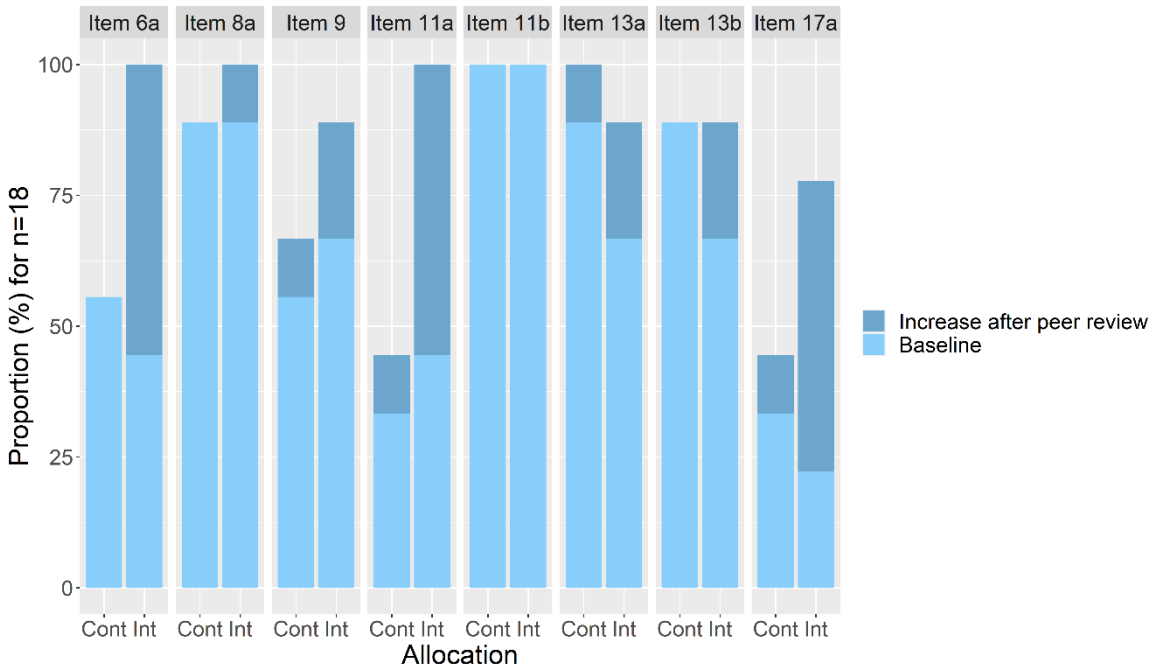


Figure 10: Proportion of manuscripts (n=18) where each CONSORT item is adequately reported. *Cont: control group; Int: intervention group. Items:*

- *6a ("Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed"),*
- *8a ("Method used to generate the random allocation sequence"), 9 ("Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned"),*
- *9 ("Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned"),*
- *11a ("If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how"),*

- *11b ("If relevant, description of the similarity of interventions"),*
- *13a ("For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome"),*
- *13b ("For each group, losses and exclusions after randomisation, together with reasons"),*
- *17a ("For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)").*

Feasibility of the intervention

The mean (SD) time taken to perform the intervention was 87 (42) minutes. Figure 11 displays a scatter plot that compares the amount of time spent to perform the intervention and the baseline score of the 12 manuscripts in the intervention group. There was no correlation between these two variables (ρ = 0.08).

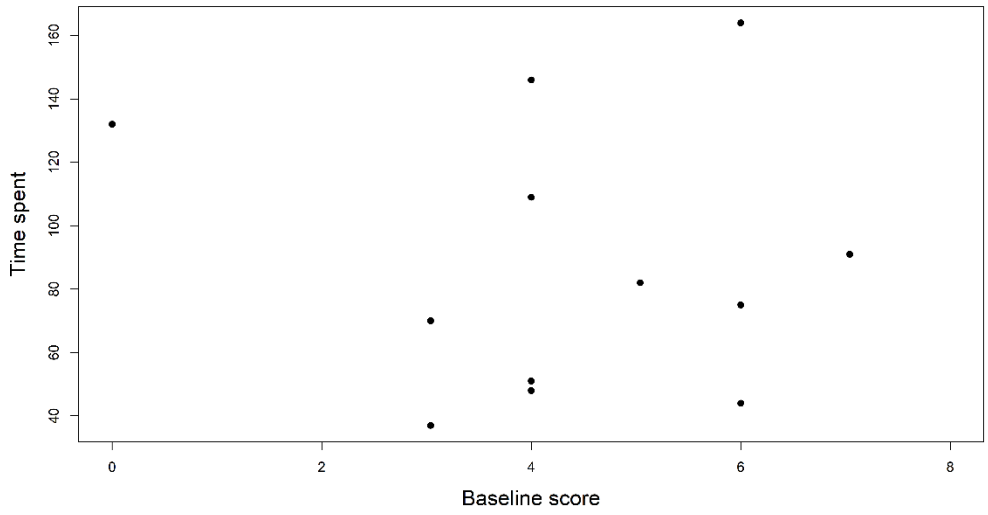


Figure 11: Scatter plot of the amount of time spent to perform the intervention and the baseline score of the 12 manuscripts in the intervention group.

Harms

We did not identify any unintended effects. There were not differences between the intervention and the control groups for (i) the proportion of manuscripts that were rejected after the first round of peer review (3 of 12, 25%, for each group). Furthermore, all authors submitted the revised manuscripts within the period requested by the handling editor.

5.4. Discussion

We found that the introduction during the peer review process of an editorial intervention performed by a researcher with expertise in the content of CONSORT significantly improved the completeness of reporting of trials submitted to BMJ Open compared to standard peer review. Six of the nine manuscripts in the intervention group achieved the maximum score and another two improved. In contrast, the only manuscript in the control group with the maximum score at follow-up already had reached that score at baseline. We observed the main differences favouring the intervention group in items 6a (Outcomes), 9 (Allocation concealment mechanism), 11a (Blinding), and 17a (Outcomes and estimation). Moreover, providing authors with extra comments on reporting issues did not seem to discourage them from revising the manuscript as all authors returned the revised manuscripts within the standard 28 days requirement.

Strengths and limitations

This study has several strengths: the randomised trial design; the fact that the intervention was performed in a real editorial context alongside peer review reports

with no disruption to usual editorial procedures; and the fact that the outcome assessment process was blinded and in duplicate.

We also note some limitations that affect the generalisability of our results. Our intervention was focused only on CONSORT, which is one of the most well-established RGs. It could potentially be more difficult for authors to fully address reviewers' comments about other less familiar RGs. We only included one journal and the same effect might not be observed in other journals. Nonetheless, we purposefully selected a very large general medical journal receiving international submissions across multiple specialties. We considered only eight core CONSORT items that are essential for evaluating the risk of bias of RCTs and not the whole checklist.

Implications

Given the importance of improving the completeness of reporting of randomised trials and given the ineffectiveness of the strategies that biomedical journals are currently implementing (14,16,144,153), it is time to take a step forward. Our study provides empirical evidence of the effectiveness of involving in the peer review process a researcher with expertise in CONSORT. In this study, the intervention was carried out by a PhD student and was implemented alongside peer review. However, this intervention could potentially be done by trained editorial staff, editors or external consultants. The demonstrated benefits of our intervention should encourage journal editors to find the best way to make this feasible.

We note that the complete-case analysis and the best-case scenario of the sensitivity analysis point to a larger effect of the intervention than the main analysis. The worst-

case scenario shows no effect. However, this scenario would assume that (1) the three rejected manuscripts in the intervention group would not improve from baseline; and that (2) all manuscripts in the control group would reach the maximum score. This scenario seems highly unlikely given that 8 out of 9 manuscripts that were not rejected in the intervention group improved from baseline and that only three controls improved and none of these reached the maximum score.

More than two decades ago, scientists started to discuss the importance of including statistical reviews as part of the publication process (180). Nowadays, statistical reviews have become widespread among top medical journals. These are usually performed by a statistician and focus on the methodological and statistical aspects of the study. As methodological issues are often not fixable, statistical reviews are key to determining the fate of manuscripts and preventing unsound research getting published (181). Completeness of reporting reviews should also become a key component in the publication system. As reporting issues are often improvable, these reviews should not generally aim to determine whether a manuscript should be published or not, but to improve their transparency. This would both help editors and peer reviewers make decisions on the manuscripts and improve the usability of published papers.

A few other RCTs have assessed different strategies for improving adherence to RGs. A recent RCT did not show that requesting authors to submit a checklist improves completeness of reporting and called for more stringent editorial policies (153). The implementation of a writing aid tool for authors (COBWEB) led to a moderate improvement in the completeness of reporting (18) whereas getting a statistician to perform an additional review against RGs showed a slightly positive but smaller than

hypothesised effect (19). Suggesting peer reviewers to check RGs (124) and implementing the web-based tool WebCONSORT at the manuscript revision stage showed no positive impact (130). However, comparisons between the results of our study and these RCTs must be made with caution as they targeted different RGs and were carried out in different settings.

The time taken for us to perform the intervention (87 minutes on average, with great variation between manuscripts) is clearly a barrier to wider implementation. Future research could evaluate whether this intervention should be focused on the whole CONSORT checklist, which would make this strategy even more time-consuming, or only on a few core items (such as those we found to be poorly reported). Also, it would be interesting to assess whether similar benefits can be obtained for other widely used RGs, such as SPIRIT (156) or PRISMA (8). Furthermore, this intervention could also be tested at other points in the editorial process, for example before the first decision is made on the manuscript or between the first decision and the invitation of external peer reviewers. For this study, we discarded both options for pragmatic reasons, as we did not want to alter the usual editorial process. While the first could be too resource intensive for journals, the latter would imply the same effort and the manuscript would undergo more transparent and accurate peer review, which could make the task of peer reviewers and handling editors easier and more efficient. We strongly recommend that journals always carry out experiments in real editorial contexts, such as this study, before considering making any changes in their policies.

Conclusions

This study provides evidence that involving a researcher with expertise in CONSORT in the process of evaluating RG checklists submitted by authors significantly improves the completeness of reporting of randomised trials. This is essential to reducing the research waste associated with inadequate reporting of RCT methods and findings. Journal editors should consider revising their peer review processes to find ways to make this intervention workable, tailoring it to their preferences.

Chapter 6: General conclusions

We have identified and explored in detail various interventions to improve authors' adherence to RGs, with a special focus on interventions related to peer review. Based on our initial scoping review (Chapter 2), which offered a big picture and helped us map the existing research on the topic of improving author adherence to RGs, we carried out a study (Chapter 3) that showed the lack of effectiveness of one of the most popular editorial interventions. Having determined that the current editorial policies are not ensuring the completeness of published research reports, we explored (Chapter 4) expert biomedical editors' opinions on what other strategies journals can follow and how to incentivise authors to comply with RGs. We tested in our final RCT (Chapter 5) the intervention that we considered most promising and feasible, and we provided empirical evidence that involving a CONSORT expert in the peer review process remarkably improves the transparency of published reports of RCTs.

Our work makes a substantial contribution within the framework of improving the efficiency of peer review and the transparency of published biomedical literature. The peer review process is often considered the "gold standard" of scholarly communication (182). However, it lacks any form of standardisation and is often biased and unable to detect important research flaws (183,184). As the development of evidence-based or clinical practice guidelines, which are essential to enhance patient care, relies on the questionable fact that published research is sound and credible, it is the duty of the scientific community to explore ways to improve peer review. Surprisingly, few interventions to improve the quality of peer review have been assessed in RCTs (185). Moreover, it is not clear what outcomes should be used to evaluate the effect of these

interventions. In the RCT reported in Chapter 5, we demonstrated an effective strategy to improve the peer review process and suggested the use of the level of transparency of the revised manuscript, measured as the degree of adherence to the corresponding RG(s), as a way to measure the impact of peer review on the quality of research reports. Despite the benefits of the intervention, we acknowledge that the time and resources journals would need to invest make it hard to widely implement the intervention. However, we propose that journals take a step forward and revise their peer review processes in order to find ways to make this intervention workable. Furthermore, our survey (Chapter 4) provides further editorial strategies and analyses their barriers and facilitators for implementation.

We also hope that our work raises awareness of the importance of transparent and accurate reporting of research. Many institutions are still using citation metrics (like the Journal Impact Factor and the H-index) to assess scientists for hiring, promotion, and tenure, what incentivises the "publish or perish" ethic. These citation metrics ignore important aspects of research quality, such as the methodological strength, the transparency of the research report, or whether researchers have followed Open Research practices (including sharing of data, protocols, software, code, materials, and other research tools) (186). For this reasons, we strongly support that publishing research completely and transparently should be used as one of the key criteria to evaluate scientists. Apart from the editorial strategies analysed in Chapters 3, 4, and 5, our initial scoping review (Chapter 2) explored further strategies that other stakeholders in the research process can follow to enhance the completeness of research reports: universities introducing RGs into graduate curricula; funders (or research ethics boards) using RGs as a template for grant (or ethics) application approvals, scientific associations

disseminating the main RGs, or research centres providing training for researchers on the use of RGs. Future experiments need to quantify the impact of these interventions before these become standard practice.

Regarding editorial interventions, we suggest that future research should investigate the effect of the intervention described in Chapter 6 in other editorial settings (different types of journals and research areas) and with other RGs that are less popular than CONSORT. Also, we propose that future studies compare the results of our RCT with the effect of other editorial strategies, such as the requirement for authors to use structured templates tailored to different study types, which is gaining popularity (141,187); the adoption of software that automatically populates RG checklists with text from the manuscript (164); or the use of reading tools that automatically assess adherence to RGs (128).

In this thesis, we have demonstrated the increase in transparency of reports of RCTs when including an expert in CONSORT in the editorial process of a biomedical journal, and we have identified and explored in detail various interventions that future research may consider evaluating. Developing and implementing effective solutions to improve adherence to RGs is a key step to increase the societal impact of biomedical research and reduce research waste.

References

1. Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, et al. Reducing waste from incomplete or unusable reports of biomedical research. Lancet. 2014 Jan 18;383(9913):267–76.

2. Ioannidis JPA, Fanelli D, Dunne DD, Goodman SN. Meta-research: Evaluation and Improvement of Research Methods and Practices. PLoS Biol. 2015 Oct;13(10):e1002264.

3. Project MiRoR. An innovative and ambitious joint doctoral training programme. [Internet]. [cited 2019 Dec 29]. Available from: http://miror-ejd.eu

4. EQUATOR Network. What is a reporting guideline? [Internet]. 2019 [cited 2019 Nov 28]. Available from: http://www.equator-network.org/about-us/what-is-a-reporting-guideline/

5. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA. 1996 Aug 28;276(8):637–9.

6. EQUATOR Network. Library for health research reporting [Internet]. 2019 [cited 2019 Dec 2]. Available from: https://www.equator-network.org/library/

7. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Bull World Health Organ. 2007;61(11):867–72.

8. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. Vol. 339, BMJ (Online). 2009. p. 332–6.

9. CONSORT group. Official extensions of CONSORT Statement [Internet]. [cited 2019 Dec 7]. Available from: http://www.consort-statement.org/extensions

10. Eldridge SM, Chan CL, Campbell MJM, Bond CM, Hopewell S, Thabane L, et al. CONSORT 2010 statement: Extension to randomised pilot and feasibility trials. BMJ. 2016;355.

11. Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. BMJ. 2012 Sep 4;345:e5661.

12. Penelope. EQUATOR/Penelope wizard [Internet]. [cited 2019 Dec 7]. Available from: https://www.penelope.ai/equator-wizard

13. Altman DG, Sauerbrei W, McShane LM. Importance of the distinction between quality of methodology and quality of reporting. Hpb. 2017;19(7):649–50.

14. Stevens A, Shamseer L, Weinstein E, Yazdi F, Turner L, Thielman J, et al. Relation of completeness of reporting of health research to journals' endorsement of reporting guidelines: systematic review. BMJ. 2014 Jun 25;348(jun25 2):g3804–g3804.

15. Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. Med J Aust. 2006 Sep 4;185(5):263–7.

16. Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. Syst Rev. 2012 Nov 29;1(1):60.

17. Samaan Z, Mbuagbaw L, Kosa D, Borg Debono V, Dillenburg R, Zhang S, et al. A systematic scoping review of adherence to reporting guidelines in health care literature. J Multidiscip Healthc. 2013 May;6:169–88.

18. Barnes C, Boutron I, Giraudeau B, Porcher R, Altman DG, Ravaud P. Impact of an online writing aid tool for writing a randomized trial report: the COBWEB (Consort-based WEB tool) randomized controlled trial. BMC Med. 2015 Sep 15;13(1):221.

19. Cobo E, Cortés J, Ribera JM, Cardellach F, Selva-O'Callaghan A, Kostov B, et al.

Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial. BMJ. 2011 Nov 22;343(nov22 2):1084.

20. Blanco D. MiRoR14 - P1 - “Scoping review on interventions to improve adherence to reporting guidelines in health research” [Internet]. [cited 2020 Apr 20]. Available from: https://zenodo.org/record/3661243

21. Blanco D. Underlying data of the project “Are CONSORT checklists submitted by authors adequately reflecting what information is actually reported in published papers?” [Internet]. [cited 2019 Dec 11]. Available from: https://zenodo.org/record/3465890

22. Blanco D. Underlying data of the project “A survey exploring biomedical editors'' perceptions of editorial interventions to improve adherence to reporting guidelines"” [Internet]. [cited 2019 Nov 29]. Available from: https://zenodo.org/record/3407725

23. Blanco D. MiRoR14 - P3 - “Effect of an editorial intervention to improve the completeness of reporting of randomised trials: a randomised controlled trial” [Internet]. [cited 2020 Apr 20]. Available from: https://zenodo.org/deposit/3661281

24. Blanco D, Altman D, Moher D, Boutron I, Kirkham JJ, Cobo E. Scoping review on interventions to improve adherence to reporting guidelines in health research. BMJ Open. 2019;9(5).

25. Blanco D, Altman DG, Moher D, Boutron I, Kirkham JJ, Cobo E. Interventions to improve adherence to reporting guidelines in health research: a scoping review protocol. BMJ Open. 2017 May 1;7(11):e017551.

26. Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc. 2015 Sep 1;13(3):141–6.

27. Arksey H, O'Malley L. Scoping studies: Towards a methodological framework. Int J Soc Res Methodol Theory Pract. 2005 Feb;8(1):19–32.

28. Mendeley Ltd. Mendeley [Internet]. [cited 2019 Dec 11]. Available from: https://www.mendeley.com/

29. Ma B, Qi G, Lin X, Wang T, Chen Z, Yang K. Epidemiology, quality, and reporting characteristics of systematic reviews of acupuncture interventions published in Chinese journals. J Altern Complement Med. 2012 Sep;18(9):813–7.

30. Larson EL, Cortazal M. Publication guidelines need widespread adoption. J Clin Epidemiol. 2012 Mar;65(3):239–46.

31. Sarkis-Onofre R, Cenci MS, Moher D, Pereira-Cenci T. Research Reporting Guidelines in Dentistry: A Survey of Editors. Braz Dent J. 2017 Feb;28(1):3–8.

32. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647–g7647.

33. Simera I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. BMC Med. 2010;8(1):24.

34. Verbeek J. Moose Consort Strobe and Miame Stard Remark or how can we improve the quality of reporting studies. Scand J Work Environ Health. 2008 Jun;34(3):165–7.

35. Kim KH, Kang JW, Lee MS, Lee JD. Assessment of the quality of reporting in randomised controlled trials of acupuncture in the Korean literature using the CONSORT statement and STRICTA guidelines. BMJ Open. 2014;4(7).

36. Moher D, Hopewell S, Schulz KF, Altman DG. Resources for authors of reports of randomized trials: Harnessing the wisdom of authors, editors, and readers. Vol. 12, Trials. 2011.

37. World Association of Medical Editors. Implementing Reporting Guidelines: Why and How, for Journal Editors [Internet]. [cited 2019 Nov 28]. Available from: https://wame.blog/2017/09/17/implementing-reporting-guidelines-why-and-how-for-journal-editors/

38. Riveros C, Dechartres A, Perrodeau E, Haneef R, Boutron I, Ravaud P. Timing and Completeness of Trial Results Posted at ClinicalTrials.gov and Published in Journals. Dickersin K, editor. PLoS Med. 2013 Dec 3;10(12):e1001566.

39. Pandis N, Turpin DL. Enhancing CONSORT compliance for improved reporting of randomized controlled trials. Vol. 145, American Journal of Orthodontics and Dentofacial Orthopedics. 2014. p. 1.

40. Rupinski M, Zagorowicz E, Regula J, Fijuth J, Kraszewska E, Polkowski M, et al. Randomized Comparison of Three Palliative Regimens Including Brachytherapy, Photodynamic Therapy and APC in Patients With Malignant Dysphagia (CONSORT 1a) (Revised II). Am J Gastroenterol. 2011 Sep 14;106(9):1612–20.

41. Kidwell CS, Liebeskind DS, Starkman S, Saver JL. Trends in acute ischemic stroke trials through the 20th century. Stroke. 2001 Jun;32(6):1349–59.

42. Sánchez-Thorin JC, Cortés MC, Montenegro M, Villate N. The quality of reporting of randomized clinical trials published in Ophthalmology. Ophthalmology. 2001;108(2):410–5.

43. Moher D, Jones A, Lepage L, CONSORT Group (Consolidated Standards for Reporting of Trials). Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. JAMA. 2001 Apr 18;285(15):1992–5.

44. Devereaux PJ, Manns BJ, Ghali WA, Quan H, Guyatt GH. The reporting of methodological factors in randomized controlled trials and the association with a journal policy to promote adherence to the Consolidated Standards of Reporting Trials (CONSORT) checklist. Control Clin Trials. 2002 Aug;23(4):380–8.

45. Montori VM, Bhandari M, Devereaux PJ, Manns BJ, Ghali WA, Guyatt GH. In the dark: the reporting of blinding status in randomized controlled trials. J Clin Epidemiol. 2002 Aug;55(8):787–90.

46. Hill CL, LaValley MP, Felson DT. Secular changes in the quality of published randomized clinical trials in rheumatology. Arthritis Rheum. 2002 Mar;46(3):779–84.

47. Faunce TA, Buckley NA. Of consents and CONSORTs: reporting ethics, law, and human rights in RCTs involving monitored overdose of healthy volunteers pre and post the "CONSORT" guidelines. J Toxicol Clin Toxicol. 2003;41(2):93–9.

48. Halpern SH, Darani R, Douglas MJ, Wight W, Yee J. Compliance with the CONSORT checklist in obstetric anaesthesia randomised controlled trials. Int J Obstet Anesth. 2004 Oct;13(4):207–14.

49. Hewitt C, Hahn S, Torgerson DJ, Watson J, Bland JM. Adequacy and reporting of allocation concealment: review of recent trials published in four general medical journals. BMJ. 2005 May 7;330(7499):1057–8.

50. Greenfield MLVH, Rosenberg AL, O'Reilly M, Shanks AM, Sliwinski MJ, Nauss MD. The Quality of Randomized Controlled Trials in Major Anesthesiology Journals. Anesth Analg. 2005 Jun;100(6):1759–64.

51. Llorca J, Martínez-Sanz F, Prieto-Salceda D, Fariñas-Alvarez C, Chinchón MV, Quinones D, et al. Quality of controlled clinical trials on glaucoma and intraocular high pressure. J Glaucoma. 2005 Jun;14(3):190–5.

52. Haahr MT, Hróbjartsson A. Who is blinded in randomized clinical trials? A study of 200 trials and a survey of authors. Clin Trials. 2006;3(4):360–5.

53. Kober T, Trelle S, Engert A. Reporting of randomized controlled trials in hodgkin lymphoma in biomedical journals. J Natl Cancer Inst. 2006 May 3;98(9):620–5.

54. Biondi-Zoccai GGL, Lotrionte M, Abbate A, Testa L, Remigi E, Burzotta F, et al.

Compliance with QUOROM and quality of reporting of overlapping meta-analyses on the role of acetylcysteine in the prevention of contrast associated nephropathy: case study. BMJ. 2006 Jan 28;332(7535):202–9.

55. Balasubramanian SP, Wiener M, Alshameeri Z, Tiruvoipati R, Elbourne D, Reed MW. Standards of reporting of randomized controlled trials in general surgery: Can we do better? Ann Surg. 2006 Nov;244(5):663–7.

56. Dias S, McNamee R, Vail A. Evidence of improving quality of reporting of randomized controlled trials in subfertility. Hum Reprod. 2006;21(10):2617–27.

57. Smidt N, Rutjes AWS, Van Der Windt DAWM, Ostelo RWJG, Bossuyt PM, Reitsma JB, et al. The quality of diagnostic accuracy studies since the STARD statement: Has it improved? Neurology. 2006 Sep;67(5):792–7.

58. Coppus SFPJ, van der Veen F, Bossuyt PMM, Mol BWJ. Quality of reporting of test accuracy studies in reproductive medicine: impact of the Standards for Reporting of Diagnostic Accuracy (STARD) initiative. Fertil Steril. 2006 Nov;86(5):1321–9.

59. Tiruvoipati R, Balasubramanian SP, Atturu G, Peek GJ, Elbourne D. Improving the quality of reporting randomized controlled trials in cardiothoracic surgery: the way forward. J Thorac Cardiovasc Surg. 2006 Aug;132(2):233–40.

60. Mahoney J, Ellison J. Assessing the quality of glucose monitor studies: a critical evaluation of published reports. Clin Chem. 2007 Jun 19;53(6):1122–8.

61. Poolman RW, Abouali JAK, Conter HJ, Bhandari M. Overlapping systematic reviews of anterior cruciate ligament reconstruction comparing hamstring autograft with bone-patellar tendon-bone autograft: why are they different? J Bone Joint Surg Am. 2007 Jul 1;89(7):1542–52.

62. Spring B, Pagoto S, Knatterud G, Kozak A, Hedeker D. Examination of the analytic quality of behavioral health randomized clinical trials. J Clin Psychol. 2007 Jan;63(1):53–71.

63. Agha R, Cooper D, Muir G. The reporting quality of randomised controlled trials

in surgery: A systematic review. Int J Surg. 2007 Dec;5(6):413–22.

64. Kane RL, Wang J, Garrard J. Reporting in randomized clinical trials improved after adoption of the CONSORT statement. J Clin Epidemiol. 2007 Mar;60(3):241–9.

65. Paranjothy B, Shunmugam M, Azuara-Blanco A. The quality of reporting of diagnostic accuracy studies in glaucoma using scanning laser polarimetry. J Glaucoma. 2007 Dec;16(8):670–5.

66. Johnson ZK, Siddiqui MAR, Azuara-Blanco A. The Quality of Reporting of Diagnostic Accuracy Studies of Optical Coherence Tomography in Glaucoma. Ophthalmology. 2007 Sep;114(9):1607–12.

67. Hind D, Booth A. Do health technology assessments comply with QUOROM diagram guidance? An empirical study. BMC Med Res Methodol. 2007 Nov 20;7:49.

68. Lee PE, Fischer HD, Rochon PA, Gill SS, Herrmann N, Bell CM, et al. Published randomized controlled trials of drug therapy for dementia often lack complete data on harm. J Clin Epidemiol. 2008 Nov;61(11):1152–60.

69. Pat K, Dooms C, Vansteenkiste J. Systematic review of symptom control and quality of life in studies on chemotherapy for advanced non-small cell lung cancer: how CONSORTed are the data? Lung Cancer. 2008 Oct;62(1):126–38.

70. Folkes A, Urquhart R, Grunfeld E. Are leading medical journals following their own policies on CONSORT reporting? Contemp Clin Trials. 2008 Nov;29(6):843–6.

71. Sinha SSS, Ashby E, Jayaram R, Grocott MPW. Quality of reporting in randomized trials published in high-quality surgical journals. J Am Coll Surg. 2009 Nov;209(5):565-571.e1.

72. Freeman K, Szczepura A, Osipenko L. Non-invasive fetal RHD genotyping tests: a systematic review of the quality of reporting of diagnostic accuracy in published studies. Eur J Obstet Gynecol Reprod Biol. 2009 Feb;142(2):91–8.

73. Uetani K, Nakayama T, Ikai H, Yonemoto N, Moher D. Quality of reports on randomized controlled trials conducted in Japan: Evaluation of adherence to the CONSORT statement. Intern Med. 2009;48(5):307–13.

74. Ethgen M, Boutron L, Steg PG, Roy C, Ravaud P. Quality of reporting internal and external validity data from randomized controlled trials evaluating stents for percutaneous coronary intervention. BMC Med Res Methodol. 2009 Apr 9;9(1):24.

75. Krzych LJ, Liszka L. No improvement in studies reporting the diagnostic accuracy of B-type natriuretic peptide. Med Sci Monit. 2009 May;15(5):SR5-14.

76. Pagoto SL, Kozak AT, John P, Bodenlos JS, Hedeker D, Spring B, et al. Intention-to-treat analyses in behavioral medicine randomized clinical trials. Int J Behav Med. 2009 Dec 25;16(4):316–22.

77. Han C, Kwak K, Marks DM, Pae C-U, Wu L-T, Bhatia KS, et al. The impact of the CONSORT statement on reporting of randomized clinical trials in psychiatry. Contemp Clin Trials. 2009 Mar;30(2):116–22.

78. Alvarez F, Meyer N, Gourraud PA, Paul C. CONSORT adoption and quality of reporting of randomized controlled trials: a systematic analysis in two dermatology journals. Br J Dermatol. 2009 Nov;161(5):1159–65.

79. Wei X, Tiejun L, Cheng W. Current situation on the reporting quality of randomized controlled trials in 5 leading Chinese medical journals. J Med Coll PLA. 2009 Apr;24(2):105–11.

80. Ladd BO, McCrady BS, Manuel JK, Campbell W. Improving the quality of reporting alcohol outcome studies: effects of the CONSORT statement. Addict Behav. 2010 Jul;35(7):660–6.

81. Yu L-M, Chan A-W, Hopewell S, Deeks JJ, Altman DG. Reporting on covariate adjustment in randomised controlled trials before and after revision of the 2001 CONSORT statement: a literature review. Trials. 2010 May 18;11(1):59.

82. Areia M, Soares M, Dinis-Ribeiro M. Quality reporting of endoscopic diagnostic studies in gastrointestinal journals: where do we stand on the use of the STARD and CONSORT statements? Endoscopy. 2010 Feb 5;42(2):138–47.

83. Delaney M, Meyer E, Cserti-Gazdewich C, Haspel RL, Lin Y, Morris A, et al. A systematic assessment of the quality of reporting for platelet transfusion studies. Transfusion. 2010 Oct;50(10):2135–44.

84. Flint HE, Harrison JE. How well do reports of clinical trials in the orthodontic literature comply with the CONSORT statement? J Orthod. 2010 Dec 16;37(4):250–61.

85. Hopewell S, Dutton S, Yu L-M, Chan A-W, Altman DG. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. BMJ. 2010 Mar 23;340(mar23 1):c723.

86. Ernst E, Hung SK CY. NCCAM-funded RCTs of herbal medicines: An independent, critical assessment. Perfusion. 2011;24:89–102.

87. Selman TJ, Morris RK, Zamora J, Khan KS. The quality of reporting of primary test accuracy studies in obstetrics and gynaecology: Application of the STARD criteria. BMC Womens Health. 2011 Mar 23;11(1):8.

88. Parsons NR, Hiskens R, Price CL, Achten J, Costa ML. A systematic survey of the quality of research reporting in general orthopaedic journals. J Bone Jt Surg - Ser B. 2011 Sep 1;93 B(9):1154–9.

89. Kiehna EN, Starke RM, Pouratian N, Dumont AS. Standards for reporting randomized controlled trials in neurosurgery. J Neurosurg. 2011 Feb;114(2):280–5.

90. Strech D, Soltmann B, Weikert B, Bauer M, Pfennig A. Quality of reporting of randomized controlled trials of pharmacologic treatment of bipolar disorders: a systematic review. J Clin Psychiatry. 2011 Sep 15;72(9):1214–21.

91. Turner L-A, Singh K, Garritty C, Tsertsvadze A, Manheimer E, Wieland LS, et al. An

evaluation of the completeness of safety reporting in reports of complementary and alternative medicine trials. BMC Complement Altern Med. 2011 Aug 22;11(1):67.

92. Haidich A-B, Birtsou C, Dardavessis T, Tirodimos I, Arvanitidou M. The quality of safety reporting in trials is still suboptimal: survey of major general medical journals. J Clin Epidemiol. 2011 Feb;64(2):124–35.

93. Gray R, Sullivan M, Altman DG, Gordon-Weeks AN. Adherence of trials of operative intervention to the CONSORT statement extension for non-pharmacological treatments: a comparative before and after study. Ann R Coll Surg Engl. 2012 Sep 1;94(6):388–94.

94. Cornelius VR, Sauzet O, Williams JE, Ayis S, Farquhar-Smith P, Ross JR, et al. Adverse event reporting in randomised controlled trials of neuropathic pain: considerations for future practice. Pain. 2013 Feb;154(2):213–20.

95. Diaz-Ordaz K, Froud R, Sheehan B, Eldridge S. A systematic review of cluster randomised trials in residential facilities for older people suggests how to improve quality. BMC Med Res Methodol. 2013 Oct 22;13(1):127.

96. Geha NN, Moseley AM, Elkins MR, Chiavegato LD, Shiwa SR, Costa LOP. The quality and reporting of randomized trials in cardiothoracic physical therapy could be substantially improved. Respir Care. 2013 Nov 1;58(11):1899–906.

97. Liu LQ, Morris PJ, Pengel LHM. Compliance to the CONSORT statement of randomized controlled trials in solid organ transplantation: a 3-year overview. Transpl Int. 2013 Mar;26(3):300–6.

98. Panic N, Leoncini E, de Belvis G, Ricciardi W, Boccia S. Evaluation of the endorsement of the preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. Derrick GE, editor. PLoS One. 2013 Dec 26;8(12):e83138.

99. Fleming PS, Seehra J, Polychronopoulou A, Fedorowicz Z, Pandis N. A PRISMA

assessment of the reporting quality of systematic reviews in orthodontics. Angle Orthod. 2013 Jan;83(1):158–63.

100. Péron J, Maillet D, Gan HK, Chen EX, You B. Adherence to CONSORT adverse event reporting guidelines in randomized clinical trials evaluating systemic cancer therapy: A systematic review. J Clin Oncol. 2013 Nov 1;31(31):3957–63.

101. Tunis AS, McInnes MDF, Hanna R, Esmail K. Association of study quality with completeness of reporting: have completeness of reporting and quality of systematic reviews and meta-analyses in major radiology journals changed since publication of the PRISMA statement? Radiology. 2013 Nov;269(2):413–26.

102. Maclean EN, Stone IS, Ceelen F, Garcia-Albeniz X, Sommer WH, Petersen SE. Quality of reporting in cardiac MRI, CT and SPECT diagnostic accuracy studies: Analysis of the impact of STARD criteria. J Cardiovasc Magn Reson. 2013 Jun 1;15(6):61.

103. Baker D, Lidster K, Sottomayor A, Amor S. Two Years Later: Journals Are Not Yet Enforcing the ARRIVE Guidelines on Reporting Standards for Pre-Clinical Animal Studies. Eisen JA, editor. PLoS Biol. 2014 Jan 7;12(1):e1001756.

104. Choi J, Jun JH, Kang BK, Kim KH, Lee MS. Endorsement for improving the quality of reports on randomized controlled trials of traditional medicine journals in Korea: a systematic review. Trials. 2014 Nov 5;15(1):429.

105. Walther S, Schueler S, Tackmann R, Schuetz GM, Schlattmann P, Dewey M. Compliance with STARD checklist among studies of coronary CT angiography: systematic review. Radiology. 2014 Apr;271(1):74–86.

106. Ghimire S, Kyung E, Lee H, Kim E. Oncology trial abstracts showed suboptimal improvement in reporting: a comparative before-and-after evaluation using CONSORT for Abstract guidelines. J Clin Epidemiol. 2014 Jun;67(6):658–66.

107. Song T-J, Leng H-F, Zhong LL, Wu T-X, Bian Z-X. CONSORT in China: past development and future direction. Trials. 2015 Jun 1;16(1):243.

108. Stevely A, Dimairo M, Todd S, Julious SA, Nicholl J, Hind D, et al. Investigation of the Shortcomings of the CONSORT 2010 Statement for the Reporting of Group Sequential Randomised Controlled Trials: A Methodological Systematic Review. Shamji M, editor. PLoS One. 2015 Nov 3;10(11):e0141104.

109. Adie S, Ma D, Harris IA, Naylor JM, Craig JC. Quality of conduct and reporting of meta-analyses of surgical interventions. Ann Surg. 2015 Apr;261(4):685–94.

110. Jull A, Aye PS. Endorsement of the CONSORT guidelines, trial registration, and the quality of reporting randomised controlled trials in leading nursing journals: A cross-sectional analysis. Int J Nurs Stud. 2015 Jun;52(6):1071–9.

111. Bearn DR, Alharbi F. Reporting of clinical trials in the orthodontic literature from 2008 to 2012: observational study of published reports in four major journals. J Orthod. 2015 Sep 15;42(3):186–91.

112. Agha RA, Fowler AJ, Limb C, Whitehurst K, Coe R, Sagoo H, et al. Impact of the mandatory implementation of reporting guidelines on reporting quality in a surgical journal: A before and after study. Int J Surg. 2016 Jun;30:169–72.

113. Pouwels KB, Widyakusuma NN, Groenwold RHH, Hak E. Quality of reporting of confounding remained suboptimal after the STROBE guideline. J Clin Epidemiol. 2016 Jan;69:217–24.

114. Rao A, Brück K, Methven S, Evans R, Stel VS, Jager KJ, et al. Quality of reporting and study design of CKD cohort studies assessing mortality in the elderly before and after STROBE: A systematic review. Zhou X, editor. PLoS One. 2016 May 11;11(5):e0155078.

115. Grob ATM, van der Vaart LR, Withagen MIJ, van der Vaart CH. Quality of reporting of diagnostic accuracy studies on pelvic floor three-dimensional transperineal ultrasound: a systematic review. Ultrasound Obstet Gynecol. 2017 Dec 21;50(4):451–7.

116. Rikos D, Dardiotis E, Tsivgoulis G, Zintzaras E, Hadjigeorgiou GM. Reporting quality

of randomized-controlled trials in multiple sclerosis from 2000 to 2015, based on CONSORT statement. Mult Scler Relat Disord. 2016 Sep;9:135–9.

117. Bigna JJR, Um LN, Nansseu JRN. A comparison of quality of abstracts of systematic reviews including meta-analysis of randomized controlled trials in high-impact general medicine journals before and after the publication of PRISMA extension for abstracts: A systematic review and m. Syst Rev. 2016 Oct 13;5(1):174.

118. Sarkis-Onofre R, Poletto-Neto V, Cenci MS, Pereira-Cenci T, Moher D. Impact of the CONSORT Statement endorsement in the completeness of reporting of randomized clinical trials in restorative dentistry. J Dent. 2017 Mar 31;58:54–9.

119. Tharyan P, Premkumar TS, Mathew V, Barnabas JP, Manuelraj. Editorial policy and the reporting of randomized controlled trials: a survey of instructions for authors and assessment of trial reports in Indian medical journals (2004-05). Natl Med J India. 21(2):62–8.

120. Lai TYY, Wong VWY, Lam RF, Cheng ACO, Lam DSC, Leung GM. Quality of reporting of key methodological items of randomized controlled trials in clinical ophthalmic journals. Ophthalmic Epidemiol. 2007;14(6):390–8.

121. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Ann Intern Med. 2015 Jun 2;162(11):777–84.

122. Jilka RL. The Road to Reproducibility in Animal Research. J Bone Miner Res. 2016 Jul;31(7):1317–9.

123. Moher D, Altman DG, Roberts I, Dirnagl U, Chalmers I, Berlin J. Four Proposals to Help Improve the Medical Research Literature. PLOS Med. 2015 Sep 22;12(9):e1001864.

124. Cobo E, Selva-O'Callagham A, Ribera JM, Cardellach F, Dominguez R, Vilardell M. Statistical reviewers improve reporting in biomedical articles: A randomized trial.

PLoS One. 2007;2(3):e332.

125. Hirst A, Altman DG. Are peer reviewers encouraged to use reporting guidelines? A survey of 116 health research journals. Cameron DW, editor. PLoS One. 2012 Apr 27;7(4):e35621.

126. Pandis N, Shamseer L, Kokich VG, Fleming PS, Moher D. Active implementation strategy of CONSORT adherence by a dental specialty journal improved randomized clinical trial reporting. J Clin Epidemiol. 2014 Sep;67(9):1044–8.

127. Hopewell S, Collins GS, Boutron I, Yu L-M, Cook J, Shanyinde M, et al. Impact of peer review on reports of randomised trials published in open peer review journals: retrospective before and after study. BMJ. 2014;349.

128. Shanahan DR. A peerless review? Automating methodological and statistical review [Internet]. [cited 2019 Nov 28]. Available from: https://blogs.biomedcentral.com/bmcblog/2016/05/23/peerless-review-automating-methodological-statistical-review/

129. Hopewell S, Ravaud P, Baron G, Boutron I. Effect of editors' implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: interrupted time series analysis. BMJ. 2012 Jun 22;344(jun22 1).

130. Hopewell S, Boutron I, Altman DG, Barbour G, Moher D, Montori V, et al. Impact of a web-based tool (WebCONSORT) to improve the reporting of randomised trials: results of a randomised controlled trial. BMC Med. 2016 Nov 28;14(1):199.

131. Mbuagbaw L, Thabane M, Vanniyasingam T, Borg Debono V, Kosa S, Zhang S, et al. Improvement in the quality of abstracts in major clinical journals since CONSORT extension for abstracts: a systematic review. Contemp Clin Trials. 2014 Jul;38(2):245–50.

132. Schriger DL, Altman DG. Inadequate post-publication review of medical research. BMJ. 2010 Aug 11;341:c3803.

133. Péron J, You B, Gan HK, Maillet D, Chen EX, Pond GR. Influence of statistician

involvement on reporting of randomized clinical trials in medical oncology. Anticancer Drugs. 2013 Mar;24(3):306–9.

134. Pandis N, Polychronopoulou A, Eliades T. An assessment of quality characteristics of randomised control trials published in dental journals. J Dent. 2010 Sep;38(9):713–21.

135. Kloukos D, Papageorgiou SN, Doulis I, Petridis H, Pandis N. Reporting quality of randomised controlled trials published in prosthodontic and implantology journals. J Oral Rehabil. 2015 Dec;42(12):914–25.

136. Goodman SN, Berlin J, Fletcher SW, Fletcher RH. Manuscript quality before and after peer review and editing at Annals of Internal Medicine. Ann Intern Med. 1994 Jul 1;121(1):11–21.

137. U.S. National Library of Medicine. ClinicalTrials.gov [Internet]. [cited 2019 Dec 11]. Available from: https://clinicaltrials.gov/

138. EQUATOR Network. Tools and templates for implementing reporting guidelines [Internet]. [cited 2019 Dec 11]. Available from: https://www.equator-network.org/toolkits/using-guidelines-in-journals/tools-and-templates-for-implementing-reporting-guidelines/

139. The PubPeer Foundation. PubPeer [Internet]. [cited 2019 Dec 11]. Available from: https://pubpeer.com/

140. ScienceOpen Inc. ScienceOpen [Internet]. [cited 2019 Dec 11]. Available from: https://www.scienceopen.com/

141. Koletsi D, Fleming PS, Behrents RG, Lynch CD, Pandis N. The use of tailored subheadings was successful in enhancing compliance with CONSORT in a dental journal. J Dent. 2017 Dec;67:66–71.

142. Cobo E, González JA. Taking advantage of unexpected WebCONSORT results. BMC Med. 2016 Dec;14(1):204.

143. Tijdink JK, Schipper K, Bouter LM, Maclaine Pont P, de Jonge J, Smulders YM. How do scientists perceive the current publication culture? A qualitative focus group interview study among Dutch biomedical researchers. BMJ Open. 2016 Feb;6(2):e008681.

144. Blanco D, Biggane AM, Cobo E. Are CONSORT checklists submitted by authors adequately reflecting what information is actually reported in published papers? Trials. 2018 Dec 29;19(1):80.

145. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: Explanation and elaboration. Ann Intern Med. 2001 Apr 17;134(8):663–94.

146. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010 Mar 23;340:c332.

147. R Core Team. R: A language and environment for statistical computing [Internet]. 2019. Available from: http://www.r-project.org/

148. Chauvin A, Moher D, Altman D, Schriger DL, Alam S, Hopewell S, et al. A protocol of a cross-sectional study evaluating an online tool for early career peer reviewers assessing reports of randomised controlled trials. BMJ Open. 2017 Sep 15;7(9):e017462.

149. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340.

150. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT Statement to Randomized Trials of Nonpharmacologic Treatment : Explanation and Elaboration. Ann Intern Med. 2008;148(4):295–309.

151. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT

statement. BMJ. 2008;337(nov11 2):a2390.

152. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009 Dec 4;339(jul21 1):b2700–b2700.

153. Hair K, Macleod MR, Sena ES. A randomised controlled trial of an Intervention to Improve Compliance with the ARRIVE guidelines (IICARus). Res Integr Peer Rev. 2019 Dec;4(1).

154. Blanco D, Hren D, Kirkham JJ, Cobo E, Schroter S. A survey exploring biomedical editors' perceptions of editorial interventions to improve adherence to reporting guidelines. F1000Research. 2019 Sep 24;8:1682.

155. Patton M. Qualitative Research & Evaluation Methods. 4th ed. Thousand Oaks: SAGE Publications; 2014.

156. Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: Defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200–7.

157. Schroter S. Evidence-based publishing [Internet]. The BMJ. [cited 2019 Nov 28]. Available from: https://www.bmj.com/about-bmj/evidence-based-publishing

158. SurveyMonkey Inc. SurveyMonkey [Internet]. 2019 [cited 2019 Nov 29]. Available from: www.surveymonkey.com

159. QSR International Pty Ltd. NVivo qualitative data analysis software. Version 12. [Internet]. 2018 [cited 2020 Apr 24]. Available from: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

160. Eysenbach G, G. E. Improving the quality of web surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. 2004 Sep 29;6(3):e34.

161. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research: A 32-item checklist for interviews and focus groups. Int J Qual Heal Care. 2018;19(6):349–57.

162. National Institute of Health. What is CME Credit? [Internet]. [cited 2019 Nov 28]. Available from: https://www.nih.gov/about-nih/what-cme-credit

163. EQUATOR Network. Toolkits [Internet]. 2019 [cited 2019 Nov 28]. Available from: https://www.equator-network.org/toolkits/

164. Hawwash D, Sharp MK, Argaw A, Kolsteren P, Lachat C. Usefulness of applying research reporting guidelines as Writing Aid software: A crossover randomised controlled trial. BMJ Open. 2019 Nov 1;9(11).

165. Publons. Global State of Peer Review 2018 [Internet]. [cited 2019 Nov 28]. Available from: https://publons.com/static/Publons-Global-State-Of-Peer-Review-2018.pdf

166. Vilaró M, Cortés J, Selva-O'Callaghan A, Urrutia A, Ribera J-M, Cardellach F, et al. Adherence to reporting guidelines increases the number of citations: the argument for including a methodologist in the editorial process and peer-review. BMC Med Res Methodol. 2019 Dec 31;19(1):112.

167. Brtnikova M, Crane LA, Allison MA, Hurley LP, Beaty BL, Kempe A. A method for achieving high response rates in national surveys of U.S. primary care physicians. Vol. 13, PLoS ONE. Public Library of Science; 2018.

168. Flanigan TS, Mcfarlane E, Cook S. Conducting Survey Research among Physicians and other Medical Professionals-A Review of Current Literature.

169. Price A, Schroter S, Clarke M, McAneney H. Role of supplementary material in biomedical journal articles: Surveys of authors, reviewers and readers. Vol. 8, BMJ Open. BMJ Publishing Group; 2018.

170. Mulligan A, Hall L, Raphael E. Peer review in a changing world: An international study measuring the attitudes of researchers. Vol. 64, Journal of the American

Society for Information Science and Technology. 2013. p. 132–61.

171. Blanco D, Schroter S, Aldcroft A, Moher D, Boutron I, Kirkham JJ, et al. Effect of an editorial intervention to improve the completeness of reporting of randomised trials: a randomised controlled trial. BMJ Open [Internet]. 2020 May 18 [cited 2020 May 19];10(5):e036799. Available from: http://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2020-036799

172. Blanco D, Schroter S, Moher D, Aldcroft A, Boutron I, Kirkham JJ, et al. Evaluating the Impact of Assessing During Peer Review the CONSORT Checklist Submitted by Authors [Internet]. 2019 [cited 2019 Nov 28]. Available from: https://clinicaltrials.gov/ct2/show/NCT03751878

173. Piaggio G, Elbourne DR, Pocock SJ, Evans SJW, Altman DG. Reporting of noninferiority and equivalence randomized trials: Extension of the CONSORT 2010 statement. Vol. 308, JAMA - Journal of the American Medical Association. 2012. p. 2594–604.

174. Vohra S, Shamseer L, Sampson M, Bukutu C, Schmid CH, Tate R, et al. CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement. J Clin Epidemiol. 2016;76:9–17.

175. Pandis N, Chung B, Scherer RW, Elbourne D, Altman DG. CONSORT 2010 statement: Extension checklist for reporting within person randomised trials. BMJ. 2017;357.

176. Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C. Reporting Randomized, Controlled Trials of Herbal Interventions: An Elaborated CONSORT Statement. Ann Intern Med. 2013 Mar 7;144(5):364.

177. MacPherson H, Altman DG, Hammerschlag R, Youping L, Taixiang W, White A, et al. Revised STandards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): Extending the CONSORT Statement. PLoS Med. 2010 Jun 8;7(6):e1000261.

178. Cheng C, Wu T, Shang H, Li Y, Altman DG, Moher D, et al. CONSORT Extension for Chinese Herbal Medicine Formulas 2017: Recommendations, Explanation, and Elaboration. Ann Intern Med. 2018;167(2):112.

179. González J-A. Randomisation of BMJ Open manuscripts [Internet]. 2019 [cited 2019 Dec 29]. Available from: http://shiny-eio.upc.edu/MIRoR/randomassign/

180. Altman DG. Statistical reviewing for medical journals. Vol. 17, Ltd. Statistics in Medicine Statist. Med. John Wiley & Sons; 1998.

181. Petrovečki M. The role of statistical reviewer in biomedical scientific journal. Biochem Medica. 2009;19(3).

182. D'Andrea R, O'Dwyer JP. Can editors save peer review from peer reviewers? PLoS One. 2017 Oct 1;12(10):e0186111.

183. Smith R. Peer review: a flawed process at the heart of science and journals. Vol. 99, J R Soc Med. 2006.

184. Stahel PF, Moore EE. Peer review for biomedical publications: We can improve the system [Internet]. Vol. 12, BMC Medicine. BioMed Central Ltd.; 2014 [cited 2020 Apr 24]. p. 179. Available from: http://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-014-0179-1

185. Bruce R, Chauvin A, Trinquart L, Ravaud P, Boutron I. Impact of interventions to improve the quality of peer review of biomedical journals: A systematic review and meta-analysis. BMC Med. 2016 Jun 10;14(1):85.

186. Moher D, Naudet F, Cristea IA, Miedema F, Ioannidis JPA, Goodman SN. Assessing scientists for hiring, promotion, and tenure. PLOS Biol. 2018 Mar 29;16(3):e2004089.

187. Trials. Structured Study Protocol Template [Internet]. [cited 2020 Jan 16]. Available from: https://trialsjournal.biomedcentral.com/submission-guidelines/preparing-your-manuscript/study-protocol/structured-study-protocol-template

Appendix A

Implementation details of the evaluated interventions identified in the scoping review.

Type of intervention	**Intervention**	**Number of studies and study design**	**Details of the intervention**	**RGs implemented**	**Format of RG implementation**	**Measure of adherence to RGs**	**Effect on adherence to RGs***
Encouraging adherence	Implementation of the writing aid tool COBWEB (18)	1 RCT	Participants have to write the six domains of the methods section of the manuscript for the protocol they receive. They have access to COBWEB tool for a random three of the six domains.	CONSORT & CONSORT extension for non-pharmacological interventions	Bullet points and examples (6 items)	Mean score for completeness of reporting (scale 0–10, items weighted)	Difference of 2.1 (95% CI 1.5-2.7)
	Author use of a structured approach for	1 Observational study (cross-sectional evaluation)	Results are posted in a standard tabular format without discussions or conclusions.	CONSORT	Checklist (4 items)	Percentage compliance of each RG item	Difference of 0.16, 0.10, 0.18 and 0.36 for each of the 4 items considered

reporting research (38)						
Journal endorsement (3 interventions, see “Details of the intervention”) (41,42,51–60,43,61–70,44,71–80,45,81–90,46,91–100,47,101–110,48,111–120,49,50)	80 observational studies (57 cross sectional evaluations of endorsing vs non-endorsing journals, 9 before and after evaluations of endorsing journals before and after endorsement, 14 both kind of evaluations)	A) Editorial statement endorsing certain RGs, B) Recommendation or requirement to follow RGs in the "Instructions to authors", and C) Requirement to submit a RG checklist together with the manuscript indicating page numbers corresponding to each item.	CONSORT (46 of 80) CONSORT extensions (9 of 80) QUOROM (3 of 80) PRISMA (4 of 80) PRISMA extensions (1 of 80) STARD (11 of 80)	Checklist (all items)	For CONSORT: percentage of compliance for each item** For other RGs: Mean summed score for completeness of reporting**	For CONSORT: 25 items improved (see details for each item on Fig. 2 on Turner et al. (16)) For CONSORT extension for harms: Difference of 0.04 (99% CI – 1.50 to 1.58) (see Stevens et al. (14)) For PRISMA: Difference of 0.53 (99% CI 0.02 to

				STROBE (4 of 80) ARRIVE (1 of 80) CONSORT, STROBE and PRISMA (11 of 80)			1.03) (see Stevens et al. (14)) For STARD: Difference of 0.52 (99% CI –0.11 to 1.16) (see Stevens et al. (14)) For STRICTA: Difference of 1.42 (99% CI –0.04 to 2.88) (see Stevens et al. (14)) For STROBE: Difference of 1.55 (99% CI –3.19 to 6.29) (see Stevens et al. (14))

	Suggestion for peer reviewers to use RGs(124)	1 RCT	Peer reviewers are sent a standard letter encouraging them to use different reporting guidelines. Reviewers are not asked to report whether they used the RG in reviewing the manuscript.	CONSORT, QUOROM, STARD	Checklist (all items)	Modified version of Manuscript Quality Assessment Instrument (scale 36-180)	Difference of 0.9 (95% CI -0.3 to +2.1)
Checking adherence and providing feedback	Completeness of reporting check by the editors (126)	1 Observational Study (Before and after evaluation)	Initial submissions are vetted by the editor-in-chief. If the submission is considered appropriate, manuscripts are assessed by the associate editor for CONSORT adherence. Authors are asked to make changes accordingly until associate editor deems appropriate that they move to the next step of the review process leading to an editorial decision.	CONSORT	Checklist (all items)	Percentage of compliance of each RG item	Before – compliance ranges from 0% to 100% (Median 40%) After – perfect compliance in 33 out of 37 items

Additional review against RGs (19)	1 RCT	A senior statistician does an additional review of all papers and provides authors suggestions on how to follow RG checklists.	STROBE, CONSORT, STARD	Checklist (all items)	Modified version of Manuscript Quality Assessment Instrument (scale 1 to 9)	Difference of 0.25 (95% CI -0.05 to +0.54)
Active implementation of RG by editors (2 interventions, see “Details of the intervention”) (129)	1 Observational study (Interrupted time series evaluation)	A) Email is sent to authors to revise the abstract according to the guidelines at the revision stage and B) Changes are made by the assistant editors of these journals towards the end of the editorial process.	CONSORT extension for abstracts	Checklist (9 of 17 items)	Monthly mean number of items reported (scale 0 to 9)	Difference of 1.5 items

	Implementation of the web-based tool WebCONSORT (130)	1 RCT	Journal editor includes a link to WebCONSORT in the revision letter to authors. Authors are directed to an automatically generated list of items and a flow diagram customised to their specific trial design.	CONSORT & some CONSORT extensions	Checklist (10 of 25 items)	Percentage of items reported for each article	Difference of 0.04 (95% CI −0.02 to +0.10)
Involvement of experts	Statistician involvement (95,133–135)	4 Observational studies (cross sectional evaluations)	Statisticians (or epidemiologists or other quantitative methodologists) are involved in the design, conduct or reporting of the study	CONSORT	Checklist (all items)	Mean score for completeness of reporting (scale 0-10, items not weighted)	In Diaz-Ordaz (95): No global effect provided (see effects for individual items in Table 2 of the paper) In Pandis et al. (133): Difference of 0.93 In Péron et al. (134): No difference in medians

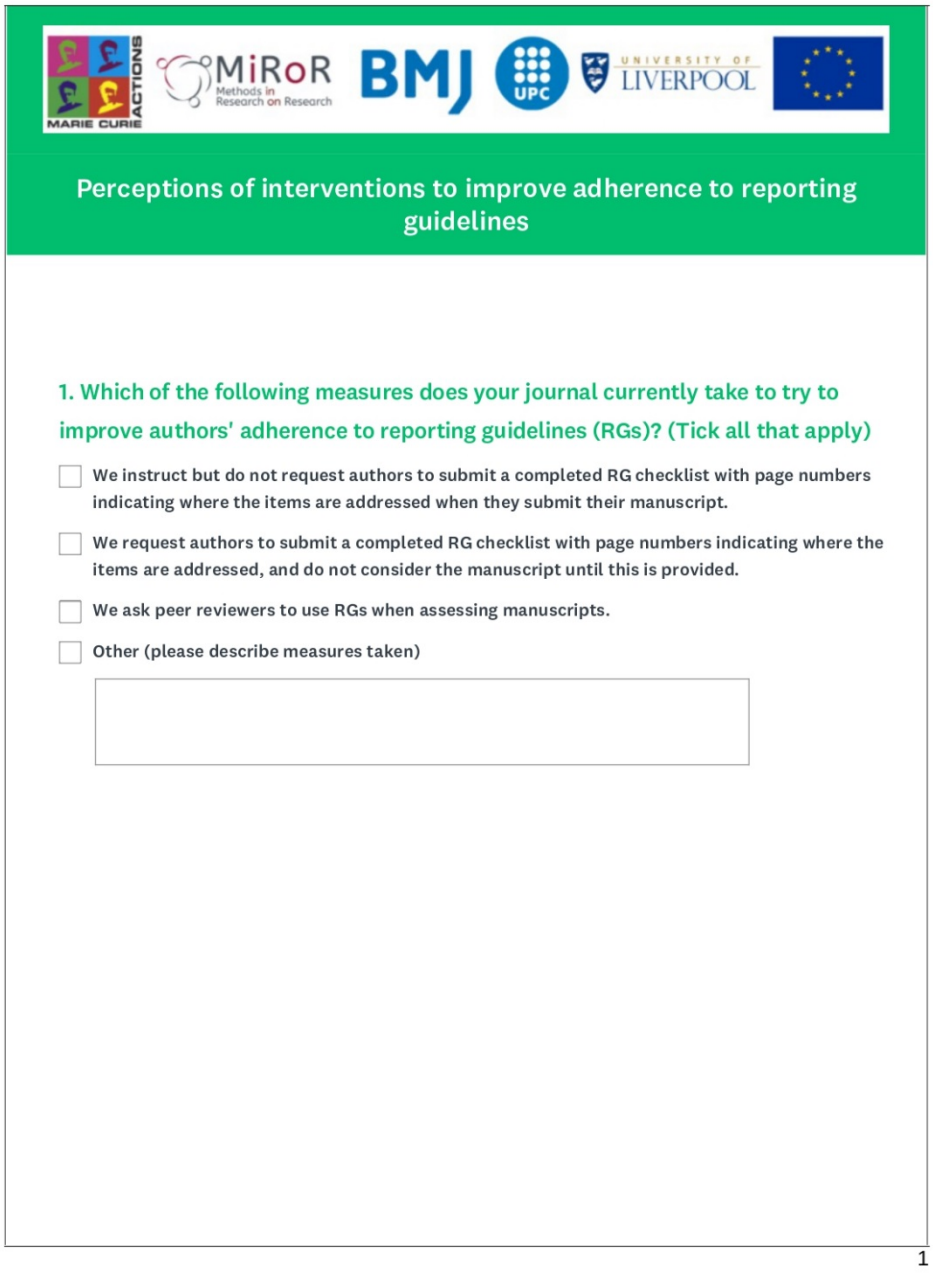
							In Kloukos et al. (135): 0.27

*Difference between adherence to RGs in intervention and non-intervention group. We did not report the CI of the effect size when authors did not report it in the original papers.

**As the 80 individual studies that belong to this category used different measures of adherence to RGs, we report here the measures used in the two systematic reviews that summarised the pooled results of most of these studies (14,16).

Appendix B

Survey questionnaire



Perceptions of interventions to improve adherence to reporting guidelines

We have identified 9 interventions to try to improve adherence to reporting guidelines that can be implemented at different points in the editorial process. For each of these interventions, we would like you to indicate how easy it would be (or was) to implement at your journal and how effective you think it would be at improving adherence to reporting guidelines if it were implemented at your journal. Moreover, you will be provided at the end of each page with an optional free text box where you can justify your answers in case you think it is needed.

2

Perceptions of interventions to improve adherence to reporting guidelines

Intervention 1 of 9: A requirement for authors to submit a completed RG checklist (using all appropriate extensions, if applicable) indicating the page numbers where each item is addressed.

Example:

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	2

2. How easy would it be (or was it) to implement this intervention at your journal?

◯ Very easy ◯ Moderately easy ◯ Neither easy nor difficult ◯ Moderately difficult

◯ Very difficult

3. If this intervention was (or is) implemented at your journal, how effective do you think it would be (or was it) at improving adherence to reporting guidelines?

◯ Very ineffective ◯ Moderately ineffective ◯ Neither ineffective nor effective

◯ Moderately effective ◯ Very effective

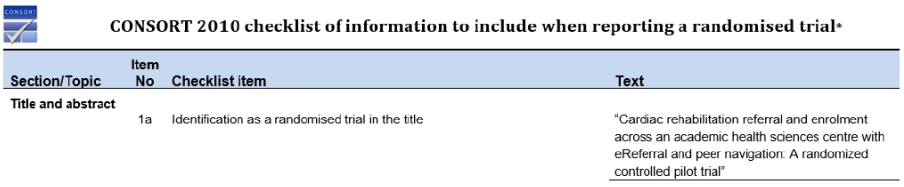
[Optional comments] You can justify your answers here:

3

Perceptions of interventions to improve adherence to reporting guidelines

Intervention 2 of 9: A requirement for authors to submit a populated RG checklist with text from their manuscript in order to facilitate the peer review process.

Example:



CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Text
Title and abstract			
	1a	Identification as a randomised trial in the title	"Cardiac rehabilitation referral and enrolment across an academic health sciences centre with eReferral and peer navigation: A randomized controlled pilot trial"

4. How easy would it be (or was it) to implement this intervention at your journal?

◯ Very easy ◯ Moderately easy ◯ Neither easy nor difficult ◯ Moderately difficult ◯ Very difficult

5. If this intervention was (or is) implemented at your journal, how effective do you think it would be (or was it) at improving adherence to reporting guidelines?

◯ Very ineffective ◯ Moderately ineffective ◯ Neither ineffective nor effective ◯ Moderately effective ◯ Very effective

[Optional comments] You can justify your answers here:

4

Perceptions of interventions to improve adherence to reporting guidelines

Intervention 3 of 9: A requirement for authors to highlight in the manuscript where each RG item is addressed.

Example:

One hundred eleven eligible patients provided informed consent to participate in the study. Data from eligible patients were collected in individual case report forms and an electronic data base was designed for the statistical analysis (CONSORT 4b).

6. How easy would it be (or was it) to implement this intervention at your journal?

◯ Very easy ◯ Moderately easy ◯ Neither easy nor difficult ◯ Moderately difficult ◯ Very difficult

7. If this intervention was (or is) implemented at your journal, how effective do you think it would be (or was it) at improving adherence to reporting guidelines?

◯ Very ineffective ◯ Moderately ineffective ◯ Neither ineffective nor effective ◯ Moderately effective ◯ Very effective

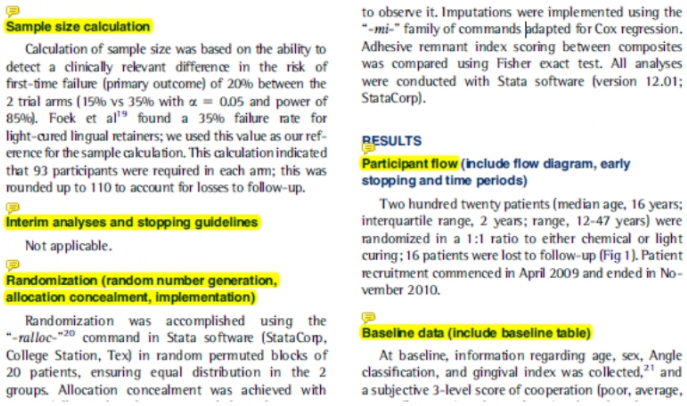
[Optional comments] You can justify your answers here:

6

Perceptions of interventions to improve adherence to reporting guidelines

Intervention 4 of 9: A requirement for authors to include new subheadings within their manuscript corresponding to different RG items within the traditional IMRaD format (Introduction, Methods, Results, and Discussion).

Example:



Sample size calculation

Calculation of sample size was based on the ability to detect a clinically relevant difference in the risk of first-time failure (primary outcome) of 20% between the 2 trial arms (15% vs 35% with $\alpha = 0.05$ and power of 85%). Foek et al[19] found a 35% failure rate for light-cured lingual retainers; we used this value as our reference for the sample calculation. This calculation indicated that 93 participants were required in each arm; this was rounded up to 110 to account for losses to follow-up.

Interim analyses and stopping guidelines

Not applicable.

Randomization (random number generation, allocation concealment, implementation)

Randomization was accomplished using the "*-ralloc-*"[20] command in Stata software (StataCorp, College Station, Tex) in random permuted blocks of 20 patients, ensuring equal distribution in the 2 groups. Allocation concealment was achieved with

to observe it. Imputations were implemented using the "*-mi-*" family of commands adapted for Cox regression. Adhesive remnant index scoring between composites was compared using Fisher exact test. All analyses were conducted with Stata software (version 12.01; StataCorp).

RESULTS

Participant flow (include flow diagram, early stopping and time periods)

Two hundred twenty patients (median age, 16 years; interquartile range, 2 years; range, 12-47 years) were randomized in a 1:1 ratio to either chemical or light curing; 16 patients were lost to follow-up (Fig 1). Patient recruitment commenced in April 2009 and ended in November 2010.

Baseline data (include baseline table)

At baseline, information regarding age, sex, Angle classification, and gingival index was collected,[21] and a subjective 3-level score of cooperation (poor, average,

8. How easy would it be (or was it) to implement this intervention at your journal?

◯ Very easy ◯ Moderately easy ◯ Neither easy nor difficult ◯ Moderately difficult

◯ Very difficult

9. If this intervention was (or is) implemented at your journal, how effective do you think it would be (or was it) at improving adherence to reporting guidelines?

◯ Very ineffective ◯ Moderately ineffective ◯ Neither ineffective nor effective

◯ Moderately effective ◯ Very effective

[Optional comments] You can justify your answers here:

8

Perceptions of interventions to improve adherence to reporting guidelines

Intervention 5 of 9: A requirement for authors on submission to use a freely available writing aid tool that guides authors through the RG checklist items, shows the key elements that need to be reported, and includes examples of adequate reporting (e.g. COBWEB).

Example:

Trial design (CONSORT 3a)

Description of trial design (such as parallel, factorial) including allocation ratio

Please describe:

- The type of trial design (parallel group, multi-arm, factorial, crossover, split-body, other)
- The conceptual framework (superiority, non-inferiority, equivalence, other)
- The allocation ratio
- Any other pertinent information (for drug development (phase 1,2,3), other)

Example. "This was a parallel-group study with imbalanced randomisation [2:1])."

9

10. How easy would it be (or was it) to implement this intervention at your journal?

◯ Very easy ◯ Moderately easy ◯ Neither easy nor difficult ◯ Moderately difficult

◯ Very difficult

11. If this intervention was (or is) implemented at your journal, how effective do you think it would be (or was it) at improving adherence to reporting guidelines?

◯ Very ineffective ◯ Moderately ineffective ◯ Neither ineffective nor effective

◯ Moderately effective ◯ Very effective

[Optional comments] You can justify your answers here:

10

Perceptions of interventions to improve adherence to reporting guidelines

Intervention 6 of 9: Instruct peer reviewers to use the appropriate RGs when assessing a manuscript.

12. How easy would it be (or was it) to implement this intervention at your journal?

◯ Very easy ◯ Moderately easy ◯ Neither easy nor difficult ◯ Moderately difficult ◯ Very difficult

13. If this intervention was (or is) implemented at your journal, how effective do you think it would be (or was it) at improving adherence to reporting guidelines?

◯ Very ineffective ◯ Moderately ineffective ◯ Neither ineffective nor effective ◯ Moderately effective ◯ Very effective

[Optional comments] You can justify your answers here:

11

Perceptions of interventions to improve adherence to reporting guidelines

Intervention 7 of 9: Instruct peer reviewers to scrutinise the completed RG checklist submitted by the authors and check its consistency with the information reported in the manuscript.

14. How easy would it be (or was it) to implement this intervention at your journal?

◯ Very easy ◯ Moderately easy ◯ Neither easy nor difficult ◯ Moderately difficult ◯ Very difficult

15. If this intervention was (or is) implemented at your journal, how effective do you think it would be (or was it) at improving adherence to reporting guidelines?

◯ Very ineffective ◯ Moderately ineffective ◯ Neither ineffective nor effective ◯ Moderately effective ◯ Very effective

[Optional comments] You can justify your answers here:

12

Perceptions of interventions to improve adherence to reporting guidelines

Intervention 8 of 9: An evaluation of the completeness of reporting by a trained editor (or editorial assistant), who would return incomplete manuscripts to authors before considering the manuscript for publication.

16. How easy would it be (or was it) to implement this intervention at your journal?

◯ Very easy ◯ Moderately easy ◯ Neither easy nor difficult ◯ Moderately difficult ◯ Very difficult

17. If this intervention was (or is) implemented at your journal, how effective do you think it would be (or was it) at improving adherence to reporting guidelines?

◯ Very ineffective ◯ Moderately ineffective ◯ Neither ineffective nor effective ◯ Moderately effective ◯ Very effective

[Optional comments] You can justify your answers here:

13

Perceptions of interventions to improve adherence to reporting guidelines

Intervention 9 of 9: Training for authors, peer reviewers, and editors on the importance, content, and use of reporting guidelines (e.g. The EQUATOR Network toolkits).

18. How easy would it be (or was it) to implement this intervention at your journal?

◯ Very easy ◯ Moderately easy ◯ Neither easy nor difficult ◯ Moderately difficult ◯ Very difficult

19. If this intervention was (or is) implemented at your journal, how effective do you think it would be (or was it) at improving adherence to reporting guidelines?

◯ Very ineffective ◯ Moderately ineffective ◯ Neither ineffective nor effective ◯ Moderately effective ◯ Very effective

[Optional comments] You can justify your answers here:

14

Perceptions of interventions to improve adherence to reporting guidelines

20. From the interventions previously described, select from this dropdown menu the one you consider to be potentially the most effective for your journal:

21. For the intervention selected, please answer the following questions:

Why do you think this intervention would be the most effective?

What would be the main difficulties in implementing this intervention?

How would you try to overcome these difficulties?

15

Perceptions of interventions to improve adherence to reporting guidelines

22. Do you have any further ideas of possible interventions to try to improve adherence to reporting guidelines? This may include modifications or combinations of the interventions previously described.

16

Perceptions of interventions to improve adherence to reporting guidelines

23. Please select which of the following best describes your current position:

- ◯ I work full time as a journal editor
- ◯ I work part time (less than 0.5 of my time) as a journal editor
- ◯ I work part time (equal or more than 0.5 of my time) as a journal editor
- ◯ Other (please specify)

24. What is your current editorial role within the journal?

- ◯ Editor in chief
- ◯ Associate editor
- ◯ Senior editor
- ◯ Other (please specify)

25. Are you involved in making decisions on the research papers and study protocols received by your journal?

- ◯ Yes
- ◯ No

17

26. For how many years have you been an editor (in total across all journals)?

- ◯ <5 years
- ◯ 5-15
- ◯ 15-25
- ◯ >25

18

Appendix C

Barriers, facilitators and possible improvements of the interventions included in the survey.

Intervention	Barriers	Facilitators	Possible improvements
I 1: A requirement for authors to submit a completed RG checklist indicating page numbers	1) Authors may overclaim adherence [Incorrect claims by authors: (i) Inconsistencies between checklist and manuscript (n=2) and (ii) N/A for applicable items (n=1)] 2) There is empirical evidence of little effectiveness in practice if compliance is not checked (14,16,144,153) (n=3) 3) Checking for compliance is resource intensive for journals (n=7)	1) Low burden on authors: quick and straightforward (n=3) 2) Easy for editors and reviewers to locate specific information (n=2)	1) Checklist to be evaluated by a trained editor or administrator (n=2)
I 2: A requirement for authors to submit a populated RG checklist with	1) Author burden – time consuming to complete checklist (n=2) 2) Checklist gets too lengthy (n=3) 3) Checking for compliance is resource intensive (n=2)	1) Manuscript length does not increase if the checklist is a supplementary file (n=2)	1) Software filling automatically the checklist (n=1)

text from their manuscript		3) Forces authors to be more rigorous (n=1)	2) Performing intervention during manuscript writing (n=1)
I 3: A requirement for authors to highlight in the manuscript where each RG item is addressed	1) Author burden – time consuming to prepare a special version of the paper (n=5) 2) Checking for compliance is resource intensive (n=4) 3) Manuscript tracking system not set up for this (n=1)	1) Easy for editors and reviewers to check adherence (n=2) 2) Everything in a single document (n=1) 3) Forces authors to be more rigorous (n=2)	1) Implement the intervention only for papers sent out to peer review (n=1)
I 4: A requirement for authors to include new subheadings	1) Author burden – time consuming to prepare a special version of the paper (n=4) 2) Requires different formats for different study designs (n=2) 3) Checking for compliance is resource intensive (n=4)	1) Easy for editors and reviewers to check adherence (n=3) 2) Easy for readers to locate information (n=1)	1) Standard templates provided to authors (n=1) [Problem: different journals having different templates (n=1)]

within their manuscript	4) Manuscript tracking system not set up for this (n=1) 5) Could cause major delays in the editorial process (n=1) 6) Maybe incompatible with journal house style (n=2) 7) May ruin the flow of the article (n=1)	3) Forces authors to respond to each item (n=2)	
I 5: A requirement for authors to use a freely available writing aid tool.	1) Author burden: (i) time consuming to rewrite the paper (n=6) 2) Checking for compliance is resource intensive (n=1) 3) Difficulty to tailor the tool to different study designs (n=2) 4) Manuscript tracking system not set up for this (n=1)	1) Better written papers (n=1) [If rejected, more chances to be accepted in the next journal (n=2)] 2) Everything is in a single document (n=1) 3) Free tool (n=1)	1) Integration of the tool in the manuscript tracking system 2) Implementation of the intervention during manuscript writing (n=1)
I 6: Instruct peer reviewers to use the appropriate	1) Peer reviewer burden: (i) too much additional work (n=2), (ii) reviewer may not know which RG to use (n=1) [Consequences:		1) Intervention should be performed by paid editorial staff (n=3)

RGs when assessing a manuscript.	effectiveness highly dependent on peer reviewers' willingness and knowledge (n=1)] 2) Difficult for editors to ensure reviewers actually use RGs (n=3)		
I 7: Instruct peer reviewers to check the consistency between the RG checklist and the manuscript.	1) Peer reviewer burden: (i) too much additional work (n=2), (ii) reviewer may not know which RG to use (n=1) [Consequences: effectiveness highly dependent on peer reviewers' willingness and knowledge (n=1)] 2) Difficult for editors to ensure reviewers actually use RGs (n=3) 3) Decrease in the quality of peer review: peer reviewers should focus on the content and not on the reporting issues (n=3)		1) Intervention should be performed by paid editorial staff (n=3)
I 8: An evaluation of the completeness of reporting by a trained editor	1) Author burden: Demoralising for authors if rejection rate is high (n=1) 2) Resource intensive for journals: (i) budget and time issues (n=9), (ii) Manuscript handling system not set up for this (n=1), (iii)	1) Authors do not have to claim adherence (n=1) 2) Credibility of the intervention (n=1)	1) Evaluation of the consistency between the manuscript and the checklist (n=1)

before the initial decision	Impractical if the volume of papers is large, (iv) Requires additional editorial staff with expertise and/or specific training (n=8) 3) In case of doubt, editors could delegate decisions (n=1)	3) Does not add more work to unpaid peer reviewers (n=2) 4) Performed by a paid and trained professional (n=13)	2) Only implement with papers sent for peer review (n=1) 3) Focus only on core items: those that enable reproducibility (n=2)
I 9: Training for authors, peer reviewers and editors on the use of RGs	1) Difficult to enforce requirement for training (n=2) 2) Resource intensive for journals: (i) budget and time issues [Especially for journals that publish a wide range of study types and dozens of RGs are needed (n=2)] 3) Requires follow-up training (n=2) 4) Editorial staff may be based in different places and time zones (n=1)	1) Intervention prior to the publication process (n=2) 2) Can target multiple stakeholders: authors, peer reviewers, and editors (n=1) 3) Can make use of existing resources on the EQUATOR website (n=1)	1) Online courses could make training more flexible (n=1) 2) Credits (162) to recognise hours of training (n=1). 3) Training could be delivered by research institutions and medical centres (n=2)