

Research Integrity in Guidelines and evidence synthesis (RIGID): a framework for assessing research integrity in guideline development and evidence synthesis

Aya Mousa,^{a,*g} Madeline Flanagan,^{b,g} Chau Thien Tay,^a Robert J. Norman,^c Michael Costello,^d Wentao Li,^{b,e} Rui Wang,^b Helena Teede,^{a,**,h} and Ben W. Mo^{b,f,h}



^aMonash Centre for Health Research and Implementation (MCHRI), Faculty of Medicine, Nursing and Health Sciences, Monash University, Victoria, Australia

^bDepartment of Obstetrics and Gynaecology, Monash University, Victoria, Australia

^cRobinson Research Institute, Adelaide Medical School, University of Adelaide, South Australia, Australia

^dUniversity of New South Wales, New South Wales, Australia

^eNational Perinatal Epidemiology and Statistics Unit, Centre for Big Data Research in Health and School of Women's and Children's Health, The University of New South Wales, Sydney, Australia

^fAberdeen Centre for Women's Health Research, University of Aberdeen, Aberdeen, UK

Summary

Background Clinical guidelines rely on sound evidence to underpin recommendations for patient care. Compromised research integrity can erode public trust and the credibility of the scientific enterprise, with potential harm to patients. Despite increased recognition of integrity concerns in scientific literature, there are no processes or guidance for incorporating integrity assessments into evidence-based guideline development or evidence synthesis more broadly.

Methods In response to this crucial gap, we developed the Research Integrity in Guidelines and evidence synthesis (RIGID) framework. Co-developed with international input from 80 multidisciplinary experts, and consumers, the RIGID framework and accompanying checklist provide an innovative and transparent six-step approach to assess the integrity of studies during the synthesis of evidence, including in the development of clinical guidelines. Central to the framework is an integrity committee, responsible for objective assessments and allocations, with constructive author engagement.

Findings The six key steps of the RIGID framework are described, as follows: (1) Review: standard systematic review processes are followed, in line with approved evidence synthesis methodologies; (2) Exclude: studies which have been retracted are excluded, and those with expressions of concern are flagged for further evaluation; (3) Assess: remaining studies are assessed for integrity using an appropriate tool and allocated an initial integrity risk rating of low, moderate or high risk for integrity concerns; (4) Discuss: integrity assessment results are discussed among integrity committee members with votes to determine final integrity risk rating allocations for each study; (5) Establish contact: low risk studies are included without author contact, whereas authors of studies ranked as moderate or high risk are contacted for clarification; (6) Reassess: studies are reassessed for inclusion using the RIGID author response algorithm (reclassified as 'included' where authors have provided a satisfactory response, 'awaiting classification' where authors have engaged but time is needed to address concerns, or 'not included' where authors have not responded to contact attempts). An illustrative case study is presented, where these six steps of the RIGID framework were successfully implemented in an influential international guideline endorsed by 39 societies across six continents. Following implementation of the framework, 45 of the 101 originally identified studies (45%) were not included in the guideline.

Interpretation Based on the latest literature and international expertise, the RIGID framework represents an important advancement in best practice standards for guideline development and evidence synthesis. Using this resource, guideline developers, policy-makers, clinicians and scientists are better positioned to navigate the currently

eClinicalMedicine
2024;74: 102717

Published Online 16 July
2024
<https://doi.org/10.1016/j.eclinm.2024.102717>

*Corresponding author. National Health and Medical Research Council (NHMRC) Research Fellow, Monash Centre for Health Research and Implementation (MCHRI), Faculty of Medicine, Nursing and Health Sciences, Monash University, 43-51 Kanooka Grove, Clayton, Victoria, 3168, Australia.

**Corresponding author. Monash Centre for Health Research and Implementation (MCHRI), NHMRC Investigator Fellow, Faculty of Medicine, Nursing and Health Sciences, Monash University, 43-51 Kanooka Grove, Clayton, Victoria, 3168, Australia.

E-mail addresses: aya.mousa@monash.edu (A. Mousa), helena.teede@monash.edu (H. Teede).

^gShared first authors.

^hShared senior authors.

precarious research landscape to ensure evidence synthesis and subsequent clinical recommendations prioritize patient care and preserve the sanctity of scientific endeavors.

Funding This study received no specific funding. The guideline in which it was piloted was supported by the Australian National Health and Medical Research Council (NHMRC) for guideline development through the Centre of Research Excellence (CRE) in Women's Health in Reproductive Life (CRE-WHiRL) (APP1171592) and the CRE in Polycystic Ovary Syndrome (CRE-PCOS) (APP1078444) led by Monash University, Australia, and partner societies: the American Society for Reproductive Medicine (ASRM), the US Endocrine Society (ENDO), the European Society of Endocrinology (ESE) and the European Society of Human Reproduction and Embryology (ESHRE).

Copyright © 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

Keywords: Research integrity; Retraction; Evidence-based guidelines; Evidence synthesis; Guideline development; Clinical practice recommendations

Research in context

Evidence before this study

Using PubMed, Scopus and Google Scholar, we conducted informal literature searches from inception up to March 2024, seeking to identify publications in English reporting on any process for incorporating research integrity into evidence synthesis or guideline development. Searches incorporated a combination of the following search terms: 'integrity', 'trustworthiness', 'research misconduct', 'retractions', 'integrity processes', 'guideline development', and 'evidence synthesis'. While tools such as the Research Integrity Assessment (RIA) and the Trustworthiness in Randomised Controlled Trials (TRACT) were identified, we found no available processes or frameworks to provide guidance on how these tools can be incorporated into guidelines or evidence synthesis processes. Given the increased recognition of integrity concerns in scientific literature, incorporating integrity assessments into evidence synthesis is critical, particularly in the context of evidence-based guidelines which directly inform patient care.

Added value of this study

In response to the pressing need to protect the interests of patients and the public, we have developed the Research Integrity in Guidelines and Evidence synthesis (RIGID) framework. Co-developed with international endorsement from 80 multidisciplinary guideline, integrity and clinical experts and consumers, this simple but innovative approach comprises six pivotal steps for incorporating integrity assessments into existing guideline and evidence synthesis processes. The RIGID framework integrates tools developed recently by Cochrane and others, which objectively assess aspects related to ethics, feasibility/plausibility and data

accuracy, irrespective of cause or intentionality. Central to the framework is the formation of an integrity committee, responsible for conducting assessments and making allocation decisions, with transparent consensus and constructive author engagement. Here, we demonstrate how the RIGID framework was successfully implemented in an influential international guideline which was led by experts in best practice guideline development and independently endorsed by 39 partner organizations across 71 countries and six continents. Using the framework, 45 of the 101 originally identified studies (45%) were not included in the guideline due to integrity concerns. Implementation of the framework is now being scaled to several international guidelines led by reputable scientific bodies.

Implications of all the available evidence

The problem of research integrity is well-recognized. Despite global efforts to address this ongoing issue, correcting the historical scientific record remains highly challenging. As we collectively grapple with the complexities of research integrity, the RIGID framework presents an important standardized process which reflects the latest literature, expert consensus, and multi-disciplinary international endorsement, and is an important advancement in best practice standards for guideline development. The framework provides sufficient flexibility to be applicable to all evidence synthesis topics and clinical questions and can be used in conjunction with existing reporting statements and quality appraisal tools. We encourage researchers, guideline developers and policy-makers to use the framework for future evidence synthesis to ensure decisions are guided by reliable and trustworthy evidence.

Introduction

Integrity in scientific research is predicated on the four pillars of honesty, accuracy, efficiency and objectivity.¹ Research integrity is particularly important in fields such as medicine, where results from clinical studies,

predominantly randomized controlled trials (RCTs), directly inform clinical guidelines, in turn shaping routine patient care. Research conducted with integrity is therefore critical to preserving public trust and maintaining the credibility of the scientific enterprise.¹

Conversely, RCTs with concerns around integrity can compromise patient care, both directly through unnecessary or harmful treatments, or indirectly through wasted resources and misdirected future medical research.² Integrity issues may arise unintentionally through honest error, incorrect analyses or naïve oversight due to inexperience, or intentionally through research misconduct. The latter encompasses data fabrication or falsification, data manipulation and/or plagiarism.³ Importantly, misconduct does not include scientific disagreements or differences in opinion.⁴

The scale of compromised research integrity within scientific research is yet to be quantified. The UK-based Committee on Publication Ethics (COPE) provides best practice guidance on managing research misconduct and, in 2009, released what is now a widely-adopted model policy for handling publication retractions.^{5,6} However, these guidelines are often insufficient, relying on authors and their institutions to respond to queries, where there is little incentive to do so. Queries to journals and authors' institutions are frequently met with silence or threats of legal action, such that official conclusions are seldom reached or formulated euphemistically, and retractions rarely result.⁷ The number of retractions in scientific literature has been estimated at 0.04%,⁸ a figure which leads some to believe that research integrity concerns are not a widespread issue. However, the median time to retraction is long, with investigations lasting up to 11.2 years in women's health for example.⁹ Given these delays and the fact that only those studies with the most obvious concerns are likely to be spotted and retracted, integrity concerns likely affect a much larger proportion of published data.⁷ In 2021, the Editor of *Anaesthesia*, analyzed individual participant data from 153 RCTs submitted from 2017 to 2020, finding that 44% contained false data and 26% had fabricated data.¹⁰ An earlier analysis by Borderwijk et al.² of 35 RCTs from one author group found that baseline characteristics in 30 of the studies (~86%) were unlikely to be the result of proper randomization. Recently, formal assessment of trustworthiness among Cochrane reviews of ante- and post-natal nutritional interventions¹¹ resulted in the removal of 25% of the RCTs due to integrity concerns. These studies affected 78% of reviews, with 72% showing a difference in effect sizes and directions, and 33% judged to require updating due to key differences in conclusions with implications for research and/or practice.¹¹

The perpetuation of problematic research is underpinned by complex systemic shortcomings, including inadequate detection systems; lack of time and resources to investigate claims; lack of incentives for journals, institutions and whistle-blowers due to concerns about reputational damage or legal implications; and most importantly, the lack of standardized procedures or protocols with appropriate oversight to manage integrity concerns. While the scientific community continues to

grapple with issues of research integrity, prompt action is urgently needed to ensure that flawed research does not extend to, nor compromise, patient care.¹²

Evidence-based guidelines are critical to guide health care and provide a vital pathway for translating research into practice. Efforts to strengthen guideline processes and improve their credibility can be traced back to the late 20th century, with the Institute of Medicine's (IOM) 1992 report, *Guidelines for Clinical Practice: From Development to Use*.¹³ A plethora of tools and standards have since emerged focusing on ensuring the integrity of the guideline development process itself. However, there is no current process to address integrity issues in the evidence informing guidelines, nor in evidence synthesis more broadly. In 2011, the IOM report, *Clinical Practice Guidelines We Can Trust*, outlined eight standards for judging the trustworthiness of guidelines, but again lacked any reference to evidence integrity.¹⁴ Similarly, the Guidelines International Network (G-I-N) resources,^{15,16} the World Health Organization's (WHO) Handbook for Guideline Development¹⁷ and the Reporting Items for practice Guidelines in Healthcare (RIGHT) Statement¹⁸ provide key principles and checklists (e.g., GIN-McMaster and CheckUp) around methodological rigor and reporting, but not on evidence integrity. The Appraisal of Guidelines for Research and Evaluation (AGREE)-II instrument evaluates the quality of clinical guidelines using six domains, including "Rigor of Development". This domain assesses the methods used to search for, select, and synthesize evidence, and evaluates the level of transparency, replicability, and accountability in the guideline development process, yet evidence integrity is still not included.¹⁹ Currently, there is no established process for assessing the integrity of studies embedded within, and informing, evidence synthesis, including in clinical guidelines.

To address this critical gap, we developed and implemented the Research Integrity in Guidelines and evidence synthesis (RIGID) framework. The aim of this framework is to provide a simplified, consistent and transparent approach for scientists, researchers, guideline developers and policy-makers to assess the integrity of the evidence and optimize the trustworthiness and reliability of subsequent recommendations for research and practice.

Methods

Framework development process

Codesign of the framework was led by an executive team with expertise in evidence synthesis, guideline development and research integrity. The framework was developed *a priori* based on broad adoption of the four-stage approach proposed by Moher and colleagues,²⁰ and subsequently piloted in the context of the International Evidence-based Guideline for Polycystic Ovary Syndrome (PCOS),^{21,22} as described below.

First, the executive team (i.e., the authors) identified the need for guidance by defining the scope of the problem² and reviewing the literature to identify existing tools and processes aimed at assessing the integrity of evidence.²³ Members of the executive team included a diverse group of guideline leads and content experts (HT, RJN, MC), methodologists (AM, CTT) as well as integrity experts (BWM, WL, RW, MF) who have published extensively on both the scale of the problem^{2,7,24–28} and the lack of tools and guidance on how to deal with this issue in evidence synthesis.²³ Having been previously aware of the issue of research integrity, it was anticipated that studies with integrity concerns would be encountered during the development of the PCOS guideline, for which there is currently no existing guidance.

The second stage involved a series of meetings and discussions aimed at developing the preliminary integrity process. Here, the executive team undertook a number of activities including developing a list of key steps for assessing integrity and identifying where and how the process would be piloted. There was general consensus among the group regarding the key elements to be included in the process. From these discussions, a six-step process for incorporating integrity was developed (as described below) and the decision was made to pilot this process within one of five guideline development groups (GDG) from the PCOS guideline. The selected GDG, focused on fertility, was where integrity issues were anticipated to be most problematic and where RCTs were primarily used, for which integrity assessment tools are available. All members of the fertility GDG were invited to attend a consensus meeting, which included members from various countries (the US, UK, Australia, Ireland, China and Vietnam) and disciplines (epidemiologists, methodologists, consumer representatives and content experts in integrity and the clinical area).

The third stage involved the meeting itself, which was held using an online, recorded forum (Zoom) among the executive group and GDG members in September 2022. The meeting started with formal presentations of the scale of the problem, empirical evidence from the literature, and the key gaps with respect to existing tools and processes. The initial planned steps and framework flow diagram were then presented to the group and feedback requested from GDG members and incorporated where applicable. All members were asked to vote for or against the proposed process, for which all voted in favor and subsequently agreed to be acknowledged in the publication of the framework. Relevant documents were circulated including the proposed framework, integrity assessment tool checklists and category descriptors, and revised timelines to incorporate this process within the fertility GDG. All members of the GDG who took part in the meeting and refinement of the process are named, with permission, in the Acknowledgements section.

Following the meeting, the framework was refined and finalized by the executive group according to feedback from the GDG. The emergent process entitled the ‘Research Integrity in Guidelines and evIDence synthesis’ or RIGID framework, was developed (Fig. 1) and piloted as described in the detailed steps and case study below. Results of the pilot were relayed back to GDG members during the guideline meetings to formulate recommendations, and the full guideline methods and recommendations, including the RIGID framework, were endorsed internationally by 80 multidisciplinary guideline, clinical, and integrity experts and consumers (listed in the publicly available guideline document²¹).

Ethics statement

This study did not involve primary data collection and therefore did not require ethics approval.

Role of the funding source

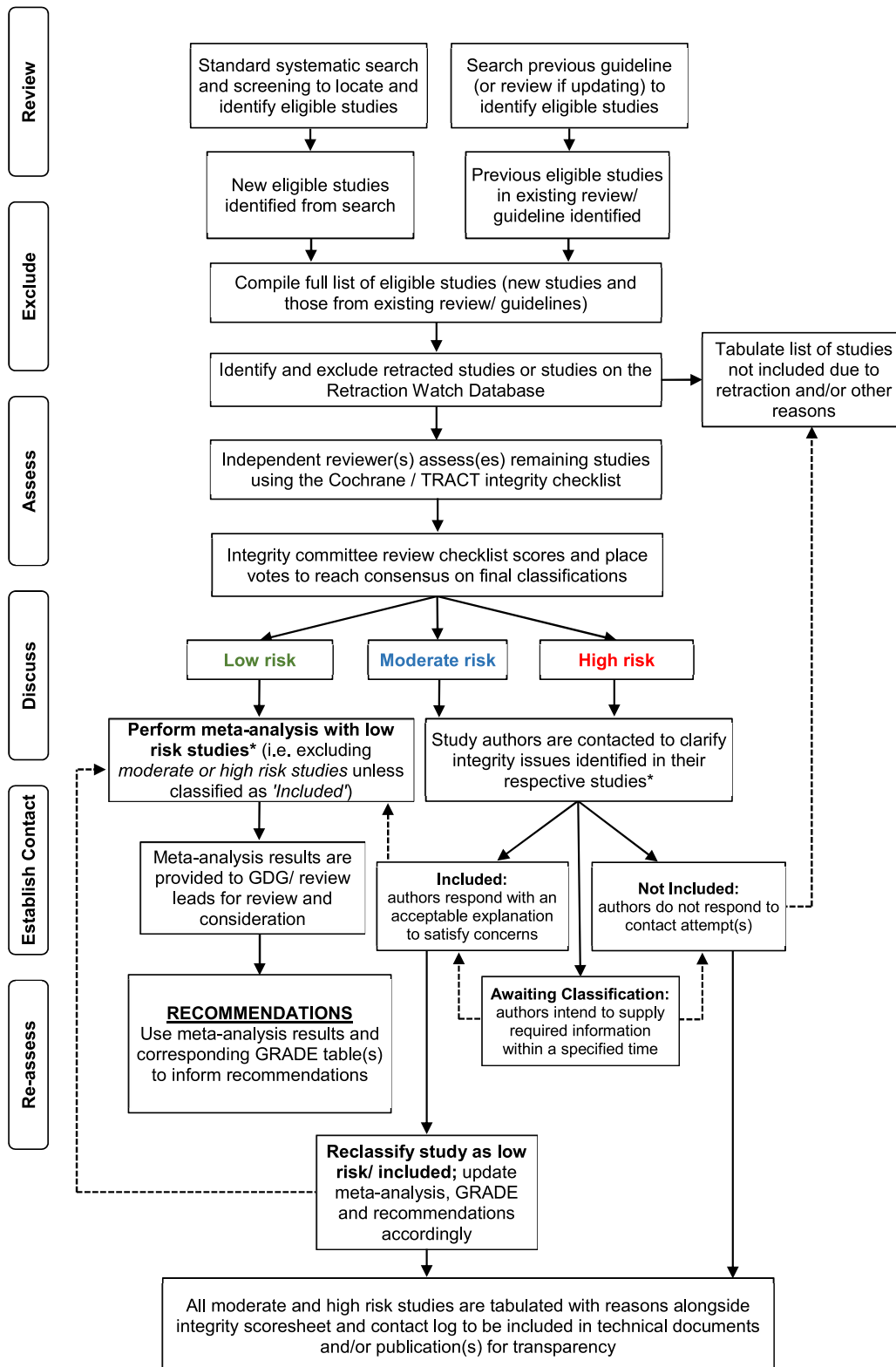
This study received no specific funding. The guideline in which the framework was piloted was supported by funders who had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

The steps described above culminated into the RIGID framework—a transparent and robust process for incorporating integrity assessments in guidelines and evidence synthesis, outlined in detail below.

Forming an integrity committee

A key step prior to undertaking the RIGID process is to establish an integrity committee, which is tasked with objective assessments regarding integrity risk ratings and final allocation of studies for the evidence synthesis. The committee is comprised of a minimum of five members, two (or more) of whom are the independent reviewers conducting the evidence review and integrity assessments (see *Step 3* below). Other members should typically represent diverse professional backgrounds both external and internal to the project or guideline, including health professionals with content expertise and/or senior clinicians (often in governance roles with an understanding of extant legislation), senior researchers with a firm understanding of the research methods under scrutiny, evidence synthesis specialists or methodologists and, if feasible, researchers with an understanding of integrity issues, which may include journal reviewers or editors. In the context of guidelines, expert panels and guideline development groups formulating the recommendations may form part, but not all, of the integrity committee to ensure diverse and external/impartial representation. Here, having a diverse, multidisciplinary group is critical to prevent both excessive and insufficient scrutiny and to ensure that decisions are made objectively, particularly for



results which may be disagreeable or contrary to prevailing beliefs. An impartial committee chair should be delegated and power dynamics should be considered when forming the committee, ensuring a balance of seniority among members of the team. All members of the committee are accountable for final decisions made regarding study allocations and are listed in subsequent publications and/or supporting documents. Hence, the decision-making process should be properly documented, transparent and defensible to ensure that the evidence synthesized is appropriately representative, and not selective.

Framework components

The RIGID framework has six essential steps, forming a rigorous, straightforward and, most importantly, transparent, method for assessing the integrity of primary research used to inform evidence synthesis. Reviewers can follow these steps without adding significant costs or delays. To simplify, we use the 'READER' acronym to summarize the six steps of the RIGID framework, outlined in [Table 1](#) and described in more detail below.

Step 1: review

The first step of the RIGID framework reflects the well-established systematic review processes already in place. Here, the reviewers follow approved methodologies for evidence synthesis and reporting (e.g., Cochrane and PRISMA^{30,31}) to search for, and identify, the relevant literature. This reflects the first steps of any systematic literature review, which have been discussed in detail elsewhere³¹ and include developing and pre-registering the protocol; establishing *a priori* the eligibility criteria, key outcomes and analyses methods; developing the search strategy; conducting the search (es); screening articles; and compiling a list of eligible studies. Once studies pass the eligibility screening stage(s) and prior to commencing data extraction, the additional components introduced by the RIGID framework, described from Step 2 below, come into effect ([Fig. 2](#)). It is important that studies with important integrity concerns are dealt with at this early stage and removed from the overall study pool (following the steps described below), not only prior to meta-analysis, but also prior to any qualitative synthesis, to ensure that conclusions are not influenced by potentially erroneous results.

Step 2: exclude

The reviewers now have a compiled list of studies which have met eligibility criteria. The first step of initiating

the integrity assessment involves identifying and excluding retracted studies. While this step may seem self-evident, there are many cases where retracted papers have been included in evidence synthesis and clinical guidelines.^{32,33} A 2022 study by Kataoka et al.³³ identified 235 systematic reviews and 17 clinical practice guidelines which cited RCTs that had already been retracted prior to their publication. Of these, 127 reviews (54%), including 11 guidelines, incorporated these already retracted RCTs into their evidence synthesis, and none corrected themselves after a median observation time of over two years.³³

By proactively searching for retraction notices or post-publication amendments issued by publishers or editors or listed on Pubpeer, and/or accessing the Retraction Watch Database (<https://retractiondatabase.org/RetractionSearch.aspx>), retracted studies can be identified and immediately excluded. Excluded studies should be tabulated with reasons for exclusion stated as 'Retracted'. Detailed guidance on how to search for retraction notices and handle retracted studies has been published by Cochrane (for Cochrane reviews, but is also relevant for other evidence synthesis processes).³⁴

Reviewers should also take note of studies with notices or expressions of concern. Unlike retractions, expressions of concern are often published by journals to raise awareness of potential problems within a published study while an investigation is underway, but prior to formal retractions. Unfortunately, these investigations are lengthy and do not align with constrained evidence synthesis or guideline development timelines, which are intended to ensure that the evidence is current. In other cases, expressions of concern are used to conclude an investigation where the outcome or evidence provided was considered inconclusive.⁶ Decisions around whether to include these studies will depend on the nature of the integrity issues (i.e., whether data validity is compromised) and should be judged on a case-by-case basis, with all decisions and reasons documented in or alongside the guideline or evidence synthesis publications. Using the RIGID framework, these studies are assessed at Step 3 below to determine whether inclusion is appropriate.

Step 3: assess

Once retracted studies are excluded, the remaining studies undergo formal assessment for research integrity. This critical step in the RIGID framework should be undertaken by two independent reviewers to limit subjectivity bias, akin to the recommended processes for

Fig. 1: Research Integrity in Guidelines and Evidence Synthesis (RIGID) Framework: a process for incorporating research integrity assessments into evidence synthesis and guideline development. GDG, guideline development group; GRADE, Grading of Recommendations, Assessment, Development and Evaluations; TRACT, Trustworthiness in Randomised Controlled Trials. *meta-analysis should not be performed until all authors have been contacted and, time-permitting, where relevant studies have been re-classified. A consistent timeframe for responses to email contact should be applied, with a suggested minimum of two weeks for initial engagement.

Phase	Description of process
1. Review	Review the literature using standard systematic review processes, in line with approved evidence synthesis methodologies (e.g., Cochrane) and compile a list of eligible studies.
2. Exclude	Exclude any studies that have been retracted or listed on the Retraction Watch Database, and note any studies that have an 'Expression of Concern' or are 'Under Investigation' by journal editors, publishers or peers (e.g., on Pubpeer).
3. Assess	Assess the integrity of the remaining studies using a well-developed tool (e.g., RIA ²⁹ or TRACT ²⁴) and allocate each study an initial integrity risk rating of low, moderate or high risk for integrity concerns.
4. Discuss	Discuss results of the integrity assessment with members of the integrity committee and place votes to determine final integrity risk rating allocations for each study.
5. Establish contact	Establish contact with authors of any studies ranked as moderate or high risk for integrity concerns to source the required information/clarification. Low risk studies do not require author contact and are included in the evidence synthesis.
6. Re-assess	Using the RIGID algorithm, re-assess studies for inclusion following a suitable timeline. Studies are categorized as 'Included' where authors have provided a satisfactory response, 'Awaiting Classification' where authors have responded with an intention to supply the requested information within a specified time; or 'Not Included' where authors have not responded to contact attempt(s). ³

RIA, Research Integrity Assessment; RIGID, Research Integrity in Guidelines and evDence synthesis; TRACT, Trustworthiness in RAnomised Controlled Trials. ³All integrity domains/judgements, votes, decisions and final rankings should be clearly reported in subsequent publications/supplemental documents of the evidence synthesis or guideline.

Table 1: Summary of READER instructions to implement the six steps of the Research Integrity in Guidelines and evDence synthesis (RIGID) framework.

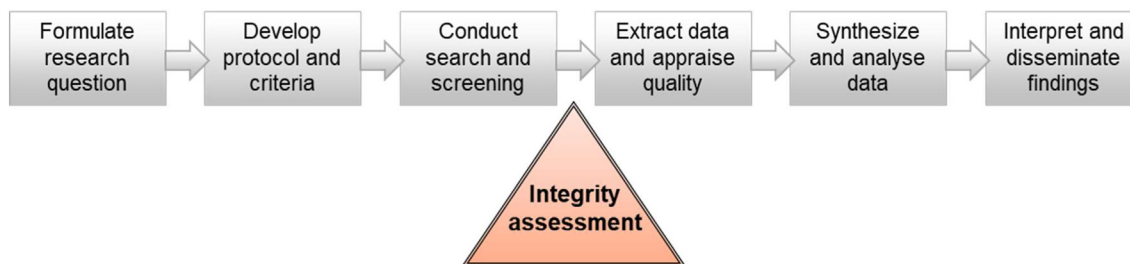


Fig. 2: Incorporating integrity assessments into existing evidence synthesis processes.

risk of bias and the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) approach.^{35,36} Independent reviewers are required to conduct assessments in a transparent manner and with justifications for ratings, to be evaluated by the integrity committee at subsequent stages.

In the last two years, a number of new assessment tools have been developed in response to the increasingly recognized issue of research integrity. In 2022, prompted by data integrity concerns for the high volume of RCTs comparing coronavirus (COVID)-19 interventions, Weibel et al.²⁹ developed the Research Integrity Assessment (RIA) tool to transparently screen RCTs and identify potential integrity issues. The RIA assesses RCTs using six study criteria which are described in detail elsewhere.²⁹ Briefly, these include study retraction, prospective trial registration, adequate ethics approval, author group, and plausibility of methods and results, with studies graded as: problematic, awaiting classification or plausible. In 2023, Mol et al.²⁴ proposed the Trustworthiness in Randomized Controlled Trials (TRACT) screening tool, which surveys seven domains including governance, author group, plausibility of intervention usage, timeframe, drop-out rates, baseline characteristics, and outcomes. Each item can be answered as either no concerns, some

concerns/no information, or major concerns, and those with several items of major concern should undergo more thorough investigation such as assessing original individual patient data.²⁴ Other tools include the Cochrane Pregnancy and Childbirth Trustworthiness Screening Tool (CPC-TST)³⁷ or the REAPPRAISED checklist (which stands for research governance, ethics, authorship, productivity, plagiarism, research conduct, analyses and methods, image manipulation, statistics and data, errors, data duplication and reporting) published in *Nature* by Grey and colleagues.³⁸

Despite minor differences, these tools are broadly designed to establish the authenticity of studies by using signaling questions to identify RCTs with potential integrity concerns. The key aim of using these tools is to pose the question 'Are the data true?'. This question has not been asked by earlier assessment tools for evidence synthesis including risk of bias,^{36,39} AMSTAR⁴⁰ or GRADE,³⁵ and demands a unique approach that cannot be adequately addressed by simply incorporating one or two additional items into these existing tools. Importantly, the integrity ratings should not be determined by comparing scores for each study, because each item individually can already lead to categorization of a study as 'high-risk' for integrity concerns (refer to published guidance for the selected tool). Use of one of these new

integrity tools is a complementary, but fundamental step, to be incorporated into well-established eligibility screening processes in evidence synthesis (Fig. 2). At the conclusion of this step, each study will have an initial integrity risk rating of either low, moderate or high risk for integrity concerns.

Step 4: discuss

With the tabulated list of studies, each with its integrity risk rating, the next step is to convene a meeting with integrity committee members to discuss and vote on study allocations. Here, the lead reviewers compile the list of eligible studies with the assessment sheet and appended publications, and these are circulated to the committee for independent review against the provided risk rating before the meeting. At the meeting, where there are discrepancies in ratings, these are discussed and areas of uncertainty or contention are raised to ensure allocated ratings are fair, then voting takes place. Ratings may shift during this process (for example, if issues initially overseen are brought to light or issues initially identified can be plausibly explained). Where a majority cannot be reached, the committee chair is tasked with the final rating decision based on the concerns presented. Vote outcomes should be documented for transparency and published in technical reports or supplemental material (see example in Table 2).

Step 5: establish contact

Studies ranked as moderate or high risk for integrity concerns are tabulated and authors are contacted by email using a standard format (see example in Appendix 1). The format should outline the guideline or evidence synthesis the study has been identified for, explaining the process and noting that some integrity issues have been identified that limit inclusion in the evidence synthesis pending clarification. Language used should be diplomatic and respectful, avoiding accusations of misconduct or reference to the credibility of researchers and, instead, focusing on the process and identified issues. It should be emphasized here that the integrity checklists utilized in RIGID aim to assess papers for the presence of unreliable data, not necessarily fabricated data. Authors are provided with the opportunity to respond and asked to indicate their willingness to engage in processes to address these issues. A consistent timeframe for responses to this email contact should be applied, with a suggested minimum of two weeks for initial engagement (ideally with a reminder after one week) and a maximum timeframe determined based on the extent of the issue(s) raised and in accordance with project timeframes. For transparency, it is critical to maintain a contact log, with detailed information on dates of contact, responses, and details of steps taken by authors to address the issues raised, as outlined in Step 6 below.

Step 6: Re-assess

There are three scenarios or pathways based on responses from authors that determine the subsequent steps to be taken, as outlined in the RIGID algorithm (Fig. 3, and more broadly in Fig. 1). Where an author response is received with additional information (e.g., clarification, data, ethics approvals or protocols) that addresses integrity issues, these studies can be moved to the 'Included' category if the studies are considered valid from an integrity perspective, and can be reliably used to inform the evidence synthesis and/or guideline recommendations. Reclassifications of studies should be considered and approved by the integrity committee.

Where there is no response by the deadline, the study moves into the 'Not Included' category (Fig. 3), based on unresolved potential integrity issues, whereby the study cannot be reliably used to inform the evidence synthesis and/or guideline recommendations. It is possible that potentially valid studies may be excluded due to lack of response, particularly when communication channels are outdated or lacking. All reasonable efforts should be made to reach authors but, once all avenues to contact authors are exhausted, it is appropriate to remove studies with no response to prevent compromising the validity and trustworthiness of the overall evidence.

In the third scenario, as occurs in many situations, the verification process requires more time (e.g., if data sharing agreements are required for original data sharing), precluding inclusion until identified integrity issues are resolved. In these cases, the study remains 'Awaiting Classification' and is not included in the current evidence synthesis (Fig. 3). While there is a risk of excluding potentially valid data, this is anticipated to be minimal with adequate time allocation. Nevertheless, prioritizing the integrity of the evidence synthesis is paramount and, as reviews and guidelines are often 'living' and undergo periodical updates, studies are re-reviewed at each update for retraction notices, published expressions of concern, or resolution of integrity issues by authors and are re-allocated on this basis. Final decisions regarding the inclusion of studies following application of the author response algorithm should be documented as shown in Table 2, and included in the publication or accompanying technical or supplemental reports.

Beyond the immediate guideline or review process, an additional optional step involves escalating concerns to journals and/or editors of published studies. Where studies with significant integrity concerns were identified and attempts to contact authors were unanswered, these studies should be flagged with the appropriate channels to initiate investigations and potentially activate a retraction process. Although not essential to inform the evidence synthesis itself, this is an important step to prevent questionable research from continuing to circulate within the scientific community and

Author, year	Governance		Author group			Plausibility of intervention			Timeframe		Drop outs		Baseline characteristics		Outcomes		Total (3a.) ^b	Initial risk rating (3b.) ^b	Voting record (4b.) ^b	Final risk rating (4b./4c.) ^b	Final study allocation (after author contact) (6a.) ^b	
	Absent or retrospective registration	Discrepant registration	Absent or vague ethics	Low # or ratio of authors	Retraction watch base (2a./2b.) ^b	Large # RCTs	Implausible intervention methods	Illogical	Fast recruitment	Fast follow-up	No LTFU	Ideal numbers	No or few (<5) BL data	Implausible data	Perfectly balanced	Larger effect size than other RCTs						Conflicting outcomes
Author, 2010	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	Retracted-Exclude	Unanimous ×6	Retracted	Excluded	
Author, 2009	Yes	No	No	Yes	No	No	No	Yes	No	No	No	Yes	No	Yes	No	Yes	No	6	High	Unanimous ×6	High	Not Included
Author, 2018	Yes	No	No	Yes	No	No	No	No	No	Yes	Yes	No	Yes	No	No	No	No	5	Moderate	×2 mod (BM, MF); ×4 high (HT, AM, JT, MC) ^a	High	Awaiting Classification
Author, 2009	No	No	No	Yes	No	Yes	No	Yes	No	No	Yes	No	Yes	Yes	No	No	No	6	High	Unanimous ×6	High	Not Included
Author, 2013	Yes	No	No	Yes	No	No	No	No	No	No	Yes	No	No	Yes	No	No	No	4	Moderate	Unanimous ×6	Moderate	Awaiting Classification
Author, 2005	No	No	No	No	Yes	Yes	No	No	No	No	Yes	No	No	No	No	No	No	3	Low	×4 low (AM, HT, JT, MC), ×2 mod (BM, MF) ^a	Low	Included
Author, 2019	Yes	No	No	Yes	No	No	No	No	No	No	Yes	No	No	Yes	No	No	No	4	Moderate	×4 low (AM, HT, JT, MC), ×2 mod (BM, MF) ^a	Low	Included

All scoring tables with votes and final decisions should be documented for transparency and published in technical reports or supplemental material. ^aDenotes the initials of committee members and their respective votes. ^bRefers to the checklist item for this part of the Research Integrity in Guidelines and Evidence synthesis (RIGID) framework (Table 3). BL, baseline; LTFU, lost to follow up; mod, moderate; TRACT, Trustworthiness in Randomized Controlled Trials tool; RCTs, randomized controlled trials. Descriptions of the domains for the tool shown here can be found in Mol et al.²⁴ and in the Supplemental Material (Appendix 2).

Table 2: Example of integrity assessment sheet to be published alongside guidelines or publications, with integrity domains assessed (using the TRACT integrity tool), risk ratings, documented votes, final risk ratings after discussion and final allocation after author contact.

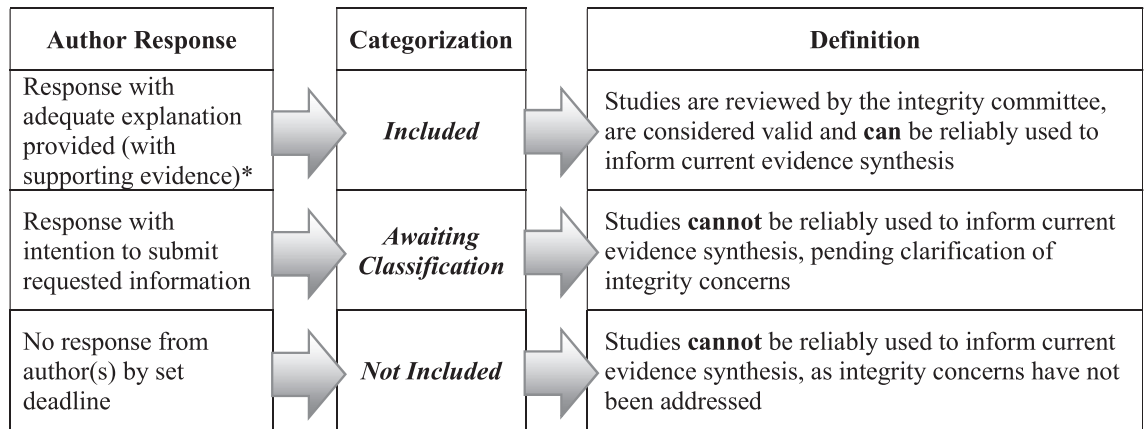


Fig. 3: Algorithm based on author(s) engagement: the final step of the Research Integrity for Guidelines and Evidence Synthesis (RIGID) framework. Author responses, or lack thereof, determine the subsequent categorization of potentially eligible studies, with inclusion determined by whether results of a study can be reliably used to inform evidence synthesis. *supporting evidence may include ethics applications, unregistered date-stamped protocols, sharing of individual participant data or other documentation as relevant, to be decided on a case-by-case basis by the integrity committee and documented transparently in publications or accompanying reports.

potentially contaminating future evidence syntheses and/or guidelines.

Case study: RIGID in action

To contextualize its utility, the RIGID framework was piloted in the 2023 International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome (PCOS).^{21,22} Governance and support were provided by the Australian National Health and Medical Research Council, and involved engagement with 39 international groups (including consumers) and 50 society nominated experts with unanimous agreement on the need, process and endorsed implementation of the framework (see case study in Fig. 4). Detailed results from this process and the groups involved are provided in the publicly available guideline²¹ and technical report.⁴¹ With its successful implementation, the framework is now being applied to other international guidelines in women’s health, including the Premature Ovarian Insufficiency (POI) international guideline and Australian adaptation of the European Society of Human Reproduction and Embryology (ESHRE) unexplained infertility guideline, for which development is currently underway.

Framework application: the RIGID checklist

A useful tool for structuring complex processes, is the use of checklists. Checklists such as the CONSORT (for randomized controlled trials),^{42,43} STROBE (for observational studies)⁴⁴ or PRISMA (for systematic reviews and meta-analyses),³⁰ among others, have become integral for the reporting of primary and secondary research, and are often a requirement for journals. Centralized resources such as the international initiative ‘Enhancing the Quality and Transparency Of health

Research’ (EQUATOR) network host useful reporting checklists aiming to support peer review and editorial decisions, and optimize reporting standards, transparency and reproducibility.⁴⁵

Informed by our pilot process, we have developed the RIGID framework checklist to enable practical integration of this approach into evidence synthesis (i.e., systematic reviews) and guideline development (Table 3). The checklist incorporates all aspects of the RIGID framework described herein, from forming an integrity committee, through to final study categorization. We recommend that researchers/reviewers and guideline developers utilize this checklist, and that overseeing organizations and peer reviewers assess the checklist in methodological review, to facilitate appropriate, consistent and transparent application of the integrity process during evidence synthesis and guideline development.

Discussion

To our knowledge, the RIGID framework is the first step-by-step process developed to incorporate research integrity assessments into evidence synthesis and guideline development. The framework offers a consistent and standardized approach, and can be integrated with existing evidence synthesis tools and processes in determining the inclusion of evidence that will inform findings and subsequent recommendations for research and practice. The process and checklist provided can, and should, be used in all forms of evidence synthesis, in conjunction with other methodological and reporting tools that relate to systematic reviews and similar review types (such as the risk of bias,^{36,39} AMSTAR,⁴⁰ and PRISMA^{30,46} series). Doing so can ensure that results are adequately reported and are

▪ **Background:** The *International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome (PCOS)* integrates the best available evidence with international, multidisciplinary clinical expertise and consumer preferences to provide healthcare professionals, consumers, educators, funders and policy makers with guidance. Leveraging substantive government and society investment, together with extensive consumer engagement and collaboration with leading experts, societies and organisations, the resulting comprehensive guideline represents a rigorous, Appraisal of Guidelines for Research and Evaluation-II (AGREE-II)-compliant guideline development process, with evidence-based recommendations to guide clinical practice.

▪ **Challenge:** Studies appearing in the guideline search results were problematic, with some retracted and others having implausible methods or results. There was no formal process for handling these studies in the guideline development process.

▪ **Solution:** The RIGID framework was developed to provide a standardised, transparent approach for incorporating integrity assessments in guideline development. An integrity committee was formed comprised of the guideline lead and one group Chair, two guideline evidence synthesis leads and two external members with research integrity expertise, totalling six members.

▪ **Governance:** The framework was supported by the Australian National Health and Medical Research Council (NHMRC) which is funding and endorsing the guideline, with broad consultation and unanimous agreement across 39 international groups (including consumers) and 50 society nominated experts on the need, process and endorsed implementation of the framework.

▪ **Implementation:** The six steps of the RIGID framework were performed as illustrated in the 'RIGID in Action' infographic on the right.

▪ **Results:** Of the original 101 studies in the pilot, a total of 45 studies (~45%) were not included in the guideline evidence synthesis, while 56 passed the integrity check and informed recommendations.

▪ **Conclusions:** The RIGID framework provided a simple and transparent process for incorporating integrity assessments in guideline development, safeguarding the integrity of the guideline itself and ensuring that recommendations for patient care are based on reliable best-practice evidence. The framework is now being implemented in another two international guidelines in infertility and premature ovarian insufficiency (POI), with further dissemination planned.

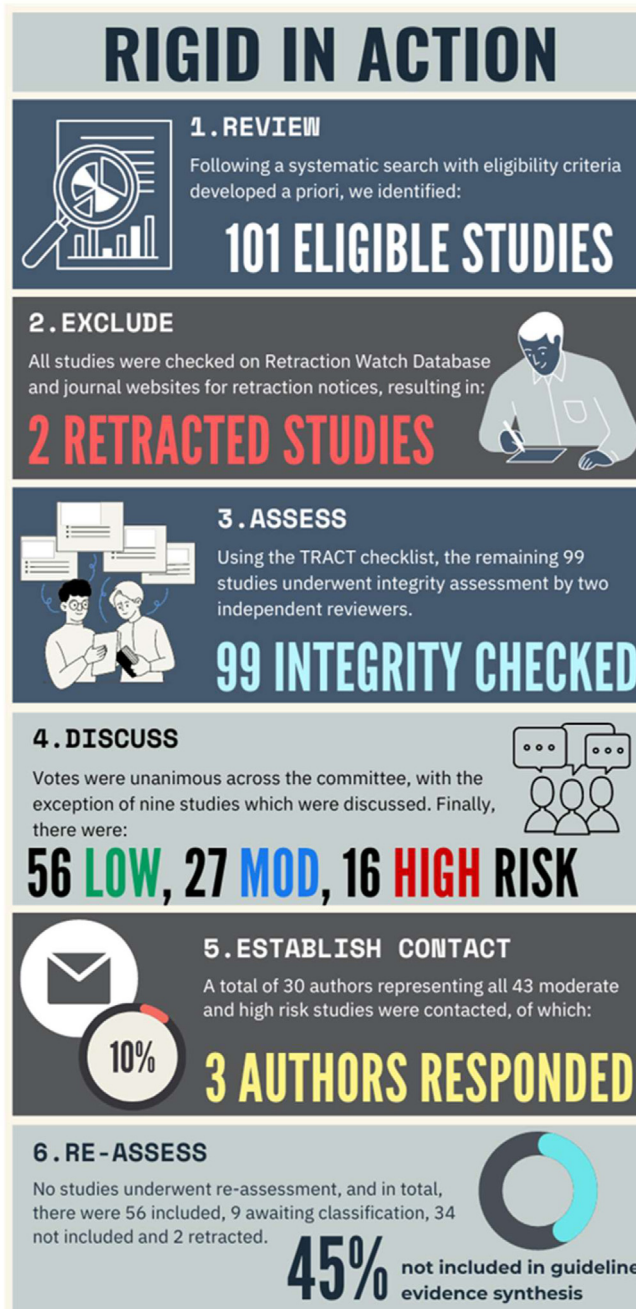


Fig. 4: Research Integrity in Guidelines and Evidence synthesis (RIGID) in Action: a case study of the RIGID framework when piloted in the context of a large international evidence-based guideline, illustrating application of the six RIGID steps. The international guideline development group who provided input and endorsed the framework are listed in the acknowledgements. MOD, moderate; RIGID, Research Integrity in Guidelines and Evidence synthesis.

Item	Description	Page
Step 0. Integrity committee		
a.	Assembled a multidisciplinary integrity committee (identified in publication(s) and/or supporting documents), comprising a minimum of five members including an impartial chair	
b.	Nominated two independent reviewers from the committee (identified in publication(s) and/or supporting documents) to conduct initial integrity assessments for each eligible study	
Step 1. Review		
1a.	For the clinical question at hand, conducted a systematic search and screening per standard review guidelines (e.g., Cochrane). <i>- This should include all steps from protocol development through to eligibility screening</i>	NA
1b.	Compiled a list of all eligible studies following full text screening	
Step 2. Exclude		
2a.	Checked all studies for retraction notices and/or on the Retraction Watch Database to identify retracted studies	
2b.	Clearly noted studies that are under investigation or have expressions of concern for further assessment	NA
2c.	All retracted studies were identified and recorded as excluded, with the reason listed as 'Retracted'	
Step 3. Assess		
3a.	Specified the tool used (e.g., TRACT or RIA) by two nominated reviewers to conduct independent integrity assessments for each study and reconcile their ratings through discussion and consensus	
3b.	Clearly documented assessments against each domain and an initial rating for each study as low, moderate or high risk of integrity concerns (with notes/justifications where relevant)	
Step 4. Discuss		
4a.	Integrity checklist assessments and risk ratings were circulated to the committee members with appended publications for review prior to the committee meeting.	NA
4b.	A meeting was convened with all committee members to discuss allocations and record votes and final risk rating after discussion. <i>Studies may be shifted from one risk rating to another following discussion</i> <i>All studies with a final rating of 'low risk' are included in the evidence synthesis</i> <i>Where a majority cannot be reached, the Chair decides the final study allocation and this is recorded, with reasons</i>	
Step 5. Establish contact		
5a.	Sourced contact details and sent a generic email to all corresponding authors of 'moderate risk' and 'high risk' studies to obtain an 'intention to respond' to concerns raised.	
5b.	Recorded a log with all authors contacted, noting those who responded (with relevant details of responses), allowing a minimum of two weeks.	
Step 6. Re-assess		
6a.	Re-assessed studies following responses (using the RIGID reassessment algorithm) and recorded final allocation as 'Included', 'Not Included' or 'Awaiting Classification'. <i>If authors are able to satisfy concerns within a reasonable timeframe, studies may be shifted to low risk and included following consultation and agreement by the integrity committee.</i>	
6b.	Continued with subsequent systematic review steps including data extraction and quality appraisal using the final list of those studies which are 'Included'	NA
<p>N.B. Any members of the integrity committee may be tasked with any of the steps listed above (e.g., the two independent reviewers may conduct integrity assessments, and a more senior member may contact authors regarding integrity concerns). For transparency, the assessment record should be provided with the guideline and/or supporting documents, outlining each domain, initial risk rating, committee votes, final risk rating after discussion, and final allocation after author contact. NA, not applicable (no page number needed for this item); RIA, research integrity assessment; RIGID, Research Integrity in Guidelines and Evidence synthesis; TRACT, Trustworthiness of Randomised Controlled Trials.</p>		
<p>Table 3: The Research Integrity in Guidelines and Evidence synthesis (RIGID) checklist.</p>		

derived from studies that are both methodologically valid and without significant integrity concerns.

It is important to recognize that the term 'integrity' is inclusive of a range of issues such as lack of ethics approval, retrospective study registration, concerns about feasibility, data analytic inaccuracies, as well as intentional fraud or fabrication of data. Integrity issues here are viewed from the perspective of reliability for clinical guidance, irrespective of their cause or intentionality. The element of intent is not the focus, and the cause(s) or reason(s) for potentially compromised integrity are secondary. Therefore, allocation of a study in the 'high risk' or 'awaiting classification' category does not equate to wrongdoing by the researchers involved. Rather, this indicates that there are matters requiring clarification, some of which may be

straightforward to explain, and studies in these categories should not be perceived any differently to those designated as high risk of bias using standard Cochrane tools.³⁶ While we acknowledge that integrity issues may adversely impact credibility, providing authors with an opportunity to respond allows for transparency and open dialogue, and should be carried out tactfully and with mutual respect.

Ultimately, the tool intends to support a research culture which upholds integrity, quality, accuracy and honesty. This reflects the purpose of publicly funded research to improve human health, and researchers are accountable to the scientific community and the general public in this regard. However, detecting integrity concerns remains highly challenging and is likely to become more problematic with the advent of artificial

intelligence. Frameworks such as RIGID must therefore be used in tandem with broader measures which enforce transparency and accountability at the individual, institutional and systemic levels. Standards and guidelines for good research practice already exist, and integrity assessments such as those proposed here are not intended to function in isolation, but rather to complement other collective efforts to maintain research quality and preserve the scientific ethos. Indeed, the last few years have seen an increased use of integrity assessment tools in general systematic reviews, including Cochrane reviews,^{11,29} but these are yet to be widely adopted or formally integrated into evidence synthesis or guideline processes. This critical gap allows potentially unreliable evidence to inform decision-making for practice, policy and resource allocation, with direct impacts on patient care. The RIGID framework addresses this gap by introducing key steps to manage integrity issues within existing evidence synthesis processes. These steps are paramount to protecting the credibility of scientific research and the welfare of its beneficiaries.

To enable consistent application and explicit consideration of integrity, we advocate for the inclusion of the RIGID framework as a core component of protocols and appraisal tools for systematic reviews and guideline development. In particular, use of the RIGID framework in the G-I-N checklists,^{15,16} WHO Handbook for Guideline Development,¹⁷ IOM Standards for Developing Trustworthy Clinical Practice Guidelines,¹⁴ and the widely used AGREE-II instrument,¹⁹ will ensure that guidelines are held to higher standards of rigor and ethical conduct. To date, best practice standards have focused on transparency and rigorous methodology through assessments of quality and risk of bias, but not of integrity *per se*. It is critical to highlight here that there are fundamental and often overlooked differences between research integrity and risk of bias or quality assessments. Integrity encompasses a broader set of principles and practices beyond methodological quality, which require explicit and transparent assessments of integrity and trustworthiness (such as TRACT) and cannot be captured through traditional tools by simply adding questions to existing risk of bias or AMSTAR evaluations. While initially, it may appear that there is overlap between integrity and other tools, they serve entirely different purposes. For instance, the ‘dropouts or loss to follow up’ criterion in the risk of bias tool by Cochrane refers to whether there is potential for bias caused by attrition (i.e., bias due to missing data). However, using an integrity checklist such as TRACT, the same ‘drop outs’ criterion refers to the plausibility of the reported losses to follow up (e.g., whether groups were perfectly balanced and if this was deemed unusual in the context of the study; or whether the study was conducted over five years without a single drop out and the plausibility of such an eventuality).

Other aspects such as recruitment timeframes, feasibility/plausibility of methods or similarities within or between tables (from other articles) are not considered in risk of bias assessments. It should also be noted that the Cochrane risk of bias tool(s)^{36,39} can categorize studies as ‘unclear’ or having ‘some concerns’ if methods are not adequately reported. This means that studies can avoid a high risk of bias rating by under-reporting; yet, continue to have an important influence on recommendations for research and clinical practice. Although some aspects of risk of bias, such as selective reporting or conflicts of interest, may indirectly touch upon integrity issues, these assessments do not adequately capture integrity, highlighting a clear need for more explicit methods. Here, we propose a simplified and standardized approach which is specific to integrity, but can be implemented into existing review processes (Fig. 2) and guideline standards (e.g., those by G-I-N,^{15,16} WHO,¹⁷ IOM¹⁴ and AGREE-II¹⁹) to directly address issues surrounding integrity in guidelines and evidence synthesis.

There are, however, some important precautions that should be taken into consideration. First, although we provide a checklist and formal process aiming to streamline what is an inherently complex, sometimes subjective and controversial decision-making process, the framework and the tools used therein (e.g., RIA, TRACT) have as yet not been validated and may result in oversimplification. Formal processes such as those described here are typically designed to capture common elements or factors relevant to a particular task or domain. However, they may not encompass all possible considerations, and important aspects may be missed or undervalued. For instance, some clinical questions may capture older trials conducted prior to the existence of trial registries or where authors are no longer contactable to clarify concerns. Others may be in countries where ethical approvals are not mandated or for which there is no formal governance. There is a risk of potentially excluding valid data originating from countries where publications occur in languages other than English (though this is an issue for evidence synthesis more broadly). While incorporating the RIGID framework is possible in these circumstances, the tools available to assess integrity may be challenging to apply, and *a priori* agreed strategies are needed, such as requiring study registration from 2010 onward, or inclusion of relevant studies in sensitivity or secondary analyses. The RIGID framework also does not directly address the issues of synthesizing secondary research (existing systematic reviews) in a guideline or umbrella review, but the principle can be applied to primary research included within these systematic reviews. To avoid overreliance on checklists and to sustain critical thinking and judgment of nuances, the RIGID framework incorporates integrated assessments by two independent reviewers and oversight and consideration by a

committee, with transparent documentation across all stages. Despite the acronym, the RIGID framework is not intended to be rigid; rather, it offers sufficient flexibility to be applied to any form of evidence synthesis whilst encouraging contextual understanding, expert judgment, and adaptability to specific circumstances. It will also likely be adapted and refined to improve relevance and effectiveness in different contexts and over time, as concerns around integrity continue to evolve.

A second key point is that, although some aspects of data integrity extend across most research contexts (e.g., peer review, replication, ethics), the procedures used to assess data integrity can vary widely between disciplines and study designs. This also applies to aspects of quality appraisal and risk of bias assessments. Methods used in qualitative and observational studies differ from those used in RCTs, with many of the current integrity tools specifically formulated for the latter, hampering standardized approaches for assessing data integrity across the research continuum. Observational studies are inherently more difficult to assess for integrity issues as they often do not require aspects such as pre-registration or well-balanced groups. This area is still in its infancy, as highlighted in the preliminary review by our group,²³ which showed that available tools for assessing misconduct in health research are elementary and labour-intensive, requiring further development, including automation and routine validation. As this active area of research gains momentum, we anticipate that further tools will become available which can be used within the RIGID framework. Meanwhile, the RIGID framework can be incorporated into RCT reviews as well as evidence-based clinical practice guidelines which rely heavily on RCT data, for which tools are currently available (e.g., RIA, TRACT, etc.).

Third, as with any tasks requiring human judgment, the application of the integrity tools and RIGID framework may be subjective or include individual bias. While allocation of two independent reviewers and an integrity committee addresses this to some extent, collective scrutiny cannot guarantee that assessments will be free from error or bias. Nevertheless, a diverse, multidisciplinary integrity committee remains key, as this brings in multiple perspectives and minimizes the likelihood of potential biases or oversights, particularly where studies present findings which challenge established norms. Collaboration with experts in research integrity, systematic review methodology, or ethics may also assist.

Fourth, introducing any process will invariably require time, effort and resources. While the RIGID framework was designed to readily integrate into existing processes, and is not intended to be time-consuming or resource-intensive, we acknowledge that challenges may arise, especially in resource-limited settings or when faced with external pressures and tight timelines. To manage this, we recommend incorporating the RIGID process from the

initial planning stages, including engaging relevant stakeholders, to ensure sufficient time and resources are allocated at the outset. On the other hand, the RIGID framework removes problematic papers early in the process, thus reducing the number of included studies and the subsequent workload needed for data extraction, analysis, quality appraisal, etc. Our experience in the PCOS guideline case study suggests that the framework can be easily integrated into established guideline processes if planned in advance. Indeed, we would argue that inadequate time is not an excuse for subpar evidence synthesis and that, the alternative, using flawed data to inform synthesis results or guideline recommendations, is not a viable option.

Finally, we need to emphasize the importance of ensuring reliability in primary data sources. While the RIGID framework focuses on assessing the reliability of papers that are already published, many of these issues in primary and subsequent secondary research could have been prevented if more attention was paid to the integrity of papers at the initial peer review.⁴⁷ Although awareness of integrity issues is growing, a more rigorous filter for assessing whether data are reliable during peer review or post-publication assessment could prevent many of these issues at the outset. Further, problematic studies identified using this framework should not be left unaddressed, as they have potential to contaminate the evidence base. While the RIGID framework focuses on aspects relevant to immediate evidence synthesis and guideline processes, we recommend that studies with integrity concerns where no response from authors is received (particularly those deemed high risk) are followed up for further investigation. Proactively raising concerns with journals and editors to potentially initiate a retraction process ensures that research integrity issues are brought to the forefront. However, we acknowledge that these are often lengthy processes which extend beyond the immediate evidence synthesis or guideline needs and, often, capacity. Overall, our pilot findings using RIGID, that 45 out of 101 studies in the international PCOS guideline could not be used due to integrity concerns, is not unique. Estimates of problematic studies varying by origin from 18% to over 90% have previously been reported.⁴⁸ If this problem is not addressed, we will ultimately need to employ more robust, but more expensive and time consuming evidence synthesis methods, including individual participant data meta-analysis.⁴⁹

In conclusion, as research practices become increasingly complex, processes are required to safeguard the community from research plagued by integrity issues. The RIGID framework presented herein provides a standardized process for incorporating integrity assessments during evidence synthesis and guideline development. The framework is not intended to be arduous or to revolutionize current methods, nor is it a standalone solution to an inherently multifaceted

problem. Rather, it provides additional, complementary steps, that can be streamlined into existing processes. These steps are critical to protecting the integrity of scientific research, and have important implications for subsequent research and practice directions. Independent of cause or intent, studies with integrity issues risk undermining public trust, and researchers, guideline developers and policy-makers have a fundamental obligation to their colleagues, to the public, and to themselves, to optimize the use of reliable evidence to inform future research efforts, resource allocation and patient care. The RIGID framework represents a significant step toward ongoing efforts to maintain the integrity of evidence synthesis and guidelines, and the evidence base upon which they are built.

Contributors

AM is the study guarantor, developed the initial framework and wrote the first draft of the manuscript, with assistance from MF. AM, MF, CTT, RN, MC, WL, RW, BM and HT contributed to conceptualization and refining of the framework and writing and editing the manuscript. All authors provided substantial intellectual input to the design, interpretation, writing and revising of the manuscript. All authors had full access to any data and/or information presented, and all were responsible for, and approved, submission of the final version for publication.

Data sharing statement

Data sharing is not applicable as no primary data were produced in this study.

Declaration of interests

All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare support from the Australian National Health and Medical Research Council (NHMRC) for guideline development through the Centre for Research Excellence (CRE) in Women's Health in Reproductive Life (CRE-WHiRL) (APP1171592) and the CRE in Polycystic Ovary Syndrome (CRE-PCOS) (APP1078444) led by Monash University, Australia, and partner societies: the American Society for Reproductive Medicine (ASRM), the US Endocrine Society (ENDO), the European Society of Endocrinology (ESE) and the European Society of Human Reproduction and Embryology (ESHRE). The primary funders of the guideline, NHMRC, were not involved in development of the guideline or the RIGID framework and have not influenced the scope. They set standards for guideline development and, based on independent peer review, approve the guideline process.

AM, RW and WL are supported by fellowships from the NHMRC. CTT receives salary support from the NHMRC-funded CRE-WHiRL and chairs their early career group, and has previously chaired the Androgen-Excess in PCOS society (AE-PCOS) early career group (2020–2022). BWM received a fellowship from the NHMRC; research funding from Merck KGaA; consulting fees for consultancy for ObsEva and Merck; and travel funding support from Merck. RN is on the Editorial committee for the Fertility and Sterility journal, chairs the committee at Westmead fertility and has a mentoring role at Flinders Fertility; and has received grant funding from the NHMRC; consulting fees from VinMec Vietnam; payment or honoraria from Cadilla Pharma for lectures; and travel support funding from Merck to speak at meetings. HT is an Executive of the International Society of Endocrinology, received a fellowship from the NHMRC, was the recipient of the NHMRC CRE-WHiRL grant and CRE-PCOS grant funding the guideline development and receives travel funding support from professional societies to attend and give plenary and symposium lectures as an independent academic and clinical expert. None of the listed organizations had any role in the design, conduct, preparation or decision to publish this manuscript.

Acknowledgements

The RIGID framework was developed by a team from the Monash University Centre for Health Research and Implementation (MCHRI) and Department of Obstetrics and Gynaecology, with international input from 80 multidisciplinary guideline, integrity, clinical and lived experience experts (full list of names and affiliations are provided in the appendices of the publicly available guideline document²¹). The implementation processes were informed by international experts from the guideline development group in the PCOS case study (see Fig. 4). With their permission, we acknowledge the members of this guideline development group: Prof. Adam Balen (Leeds Teaching Hospitals and British Fertility Society, UK); Dr Lisa Bedson (Repromed, Australia); Prof. Roger Hart (University of Western Australia, Australia); Dr Tuong Ho (HOPE Research Centre, My Duc Hospital, Vietnam); Dr Kim Hopkins (PCOS Challenge: The National Polycystic Ovary Syndrome Association, USA); Cailin Jordan (Genea Hollywood Fertility, Australia); Prof. Richard Legro (Penn State Clinical and Translational Institute, USA); Dr Edgar Mocuana (Rotunda Hospital, Ireland); Prof. Luk Rombaouts (Monash University, Australia); Prof. Shakila Thangaratnam (University of Birmingham, NIHR Biomedical Research Centre, Birmingham, UK); Prof. Dongzi Yang (Reproductive Medical Centre, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, China).

No members of the public or patients were directly involved in preparing the manuscript. As noted above, the framework was endorsed by an international guideline development group (GDG), which included a consumer representative from a patient advocacy group. The consumer representative is listed above (with permission) with the other members of the GDG, all of whom provided feedback on the framework in the context of the international guideline implementation pilot (i.e., the case study).

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.eclim.2024.102717>.

References

- 1 Steneck NH. *Office of research integrity (ORI): introduction to the responsible conduct of research*. USA: U.S. Department of Health and Human Services; 2007.
- 2 Bordewijk EM, Wang R, Askie LM, et al. Data integrity of 35 randomised controlled trials in women's health. *Eur J Obstet Gynecol Reprod Biol.* 2020;249:72–83.
- 3 Robishaw JD, DeMets DL, Wood SK, Boisselle PM, Hennekens CH. Establishing and maintaining research integrity at academic institutions: challenges and opportunities. *Am J Med.* 2020;133(3):e87–e90.
- 4 Resnik DB, Stewart CN Jr. Misconduct versus honest error and scientific disagreement. *Account Res.* 2012;19(1):56–63.
- 5 COPE. COPE Promoting integrity in research and its publication. <https://publicationethics.org/>; 2023. Accessed April 9, 2023.
- 6 COPE Council. *COPE retraction guidelines - English*. Committee on Publication Ethics; 2019. Version 2: November 2019.
- 7 Li W, Gurrin LC, Mol BW. Violation of research integrity principles occurs more often than we think. *Reprod Biomed Online.* 2022;44(2):207–209.
- 8 Brainard JYJ. What a massive database of retracted papers reveals about science publishing's 'death penalty'. *Science.* 2018;25:1–5.
- 9 Li W, Bordewijk EM, Mol BW. Assessing research misconduct in randomized controlled trials. *Obstet Gynecol.* 2021;138(3):338–347.
- 10 Carlisle JB. False individual patient data and zombie randomised controlled trials submitted to Anaesthesia. *Anaesthesia.* 2021;76(4):472–479.
- 11 Weeks J, Cuthbert A, Alfirevic Z. Trustworthiness assessment as an inclusion criterion for systematic reviews—What is the impact on results? *Cochrane Ev Synth.* 2023:e12037. <https://doi.org/10.1002/cesm.12037>.
- 12 Abbasi K. Retract or be damned: a dangerous moment for science and the public. *BMJ.* 2023;381:p1424.
- 13 Institute of Medicine Committee on Clinical Practice G, Field MJ, Lohr KN. *Guidelines for clinical practice: from development to use*. Washington (DC): National Academies Press (US) Copyright 1992 by the National Academy of Sciences; 1992.
- 14 IOM. *Clinical practice guidelines we can trust*. Washington, DC: The National Academies Press; 2011.

- 15 Vernooij RWM, Alonso-Coello P, Brouwers M, Martínez García L, CheckUp P. Reporting items for updated clinical guidelines: checklist for the reporting of updated guidelines (CheckUp). *PLoS Med.* 2017;14(1):e1002207.
- 16 Guidelines International Network. GIN guidelines international network resources. <https://g-i-n.net/get-involved/resources>.
- 17 World Health O. *WHO handbook for guideline development*. 2nd ed. Geneva: World Health Organization; 2014.
- 18 Chen Y, Yang K, Marušić A, et al. A reporting tool for practice guidelines in health care: the RIGHT statement. *Ann Intern Med.* 2017;166(2):128–132.
- 19 Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ.* 2010;182(18):E839–E842.
- 20 Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. *PLoS Med.* 2010;7(2):e1000217.
- 21 Teede H, Tay CT, Laven JSE, et al. International evidence-based guideline for the assessment and management of polycystic ovary syndrome 2023. https://bridges.monash.edu/articles/online_resource/International_Evidence-based_Guideline_for_the_Assessment_and_Management_of_Polycystic_Ovary_Syndrome_2023/24003834;2023.
- 22 Teede HJ, Tay CT, Laven JJE, et al. Recommendations from the 2023 international evidence-based guideline for the assessment and management of polycystic ovary Syndrome. *J Clin Endocrinol Metab.* 2023;108(10):2447–2469.
- 23 Bordewijk EM, Li W, van Eekelen R, et al. Methods to assess research misconduct in health-related research: a scoping review. *J Clin Epidemiol.* 2021;136:189–202.
- 24 Mol BW, Lai S, Rahim A, et al. Checklist to assess Trustworthiness in RANdomised Controlled Trials (TRACT checklist): concept proposal and pilot. *Res Integr Peer Rev.* 2023;8(1):6.
- 25 Linn MM, Mol BW. Data integrity assessment in obstetrics and gynaecology. *J Gynecol Obstet Hum Reprod.* 2022;51(8):102440.
- 26 Li W, van Wely M, Gurrin L, Mol BW. Integrity of randomized controlled trials: challenges and solutions. *Fertil Steril.* 2020;113(6):1113–1119.
- 27 Bordewijk EM, Li W, Gurrin LC, Thornton JG, van Wely M, Mol BW. An investigation of seven other publications by the first author of a retracted paper due to doubts about data integrity. *Eur J Obstet Gynecol Reprod Biol.* 2021;261:236–241.
- 28 Liu Y, Thornton JG, Li W, van Wely M, Mol BW. Concerns about data integrity of 22 randomized controlled trials in women's health. *Am J Perinatol.* 2021;40(3):279–289.
- 29 Weibel S, Popp M, Reis S, Skoetz N, Garner P, Sydenham E. Identifying and managing problematic trials: a research integrity assessment tool for randomized controlled trials in evidence synthesis. *Res Synth Methods.* 2022;14(3):357–369. n/a(n/a).
- 30 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ.* 2009;339:b2535.
- 31 Higgins JPTJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA. *Cochrane Handbook for Systematic Reviews of Interventions version 6.4*. www.training.cochrane.org/handbook; 2023.
- 32 Bolland MJ, Grey A, Avenell A. Citation of retracted publications: a challenging problem. *Account Res.* 2022;29(1):18–25.
- 33 Kataoka Y, Banno M, Tsujimoto Y, et al. Retracted randomized controlled trials were cited and not corrected in systematic reviews and clinical practice guidelines. *J Clin Epidemiol.* 2022;150:90–97.
- 34 Lefebvre C, Glanville J, Briscoe S, et al. On behalf of the cochrane information retrieval methods group. Technical supplement to chapter 4: searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for systematic reviews of interventions version 6.3*. Cochrane; 2022. www.training.cochrane.org/handbook.
- 35 Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ.* 2008;336(7650):924.
- 36 Higgins JPT, Altman DG, Gøtzsche PC, et al. The cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ.* 2011;343:d5928.
- 37 Alfirevic Z, Kellie FJ, Stewart F, Jones L, Hampson L. On behalf of pregnancy and Childbirth, editorial board. Identifying and handling potentially untrustworthy trials in pregnancy and childbirth cochrane reviews. https://pregnancy.cochrane.org/sites/pregnancy.cochrane.org/files/public/uploads/identifying_and_handling_potentially_untrustworthy_trials_v.2.4_-_20_july_2021.pdf; 2021.
- 38 Grey A, Bolland MJ, Avenell A, Klein AA, Gunsalus C. Check for publication integrity before misconduct. *Nature.* 2020;577(7789):167–169.
- 39 Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ.* 2019;366:l4898.
- 40 Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ.* 2017;358:j4008.
- 41 Mousa A, Tay CT, Teede H. *Technical report for the 2023 international evidence-based guideline for the assessment and management of polycystic ovary syndrome*. Monash University; 2023.
- 42 Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA.* 1996;276(8):637–639.
- 43 Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. *BMC Med Res Methodol.* 2001;1:2.
- 44 von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ.* 2007;335(7624):806–808.
- 45 Han S, Olonisakin TF, Pribis JP, et al. A checklist is associated with increased quality of reporting preclinical biomedical research: a systematic review. *PLoS One.* 2017;12(9):e0183591.
- 46 Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev.* 2021;10(1):89.
- 47 Wang R, van Wely M. Understand low-quality evidence: learn from food chains. *Fertil Steril.* 2020;113(1):93–94.
- 48 Ioannidis JPA. Hundreds of thousands of zombie randomised trials circulate among us. *Anaesthesia.* 2021;76(4):444–447.
- 49 Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ.* 2010;340:c221.