Article:
Faltinsen, Erlend, Østengaard, Lasse, Boutron, Isabelle et al. (2025) User experience of Tool for Addressing Conflicts of Interest in Trials (TACIT) prototype: interview and questionnaire study. Journal of Clinical Epidemiology. 112003. ISSN: 0895-4356

https://doi.org/10.1016/j.jclinepi.2025.112003

Journal of Clinical Epidemiology 188 (2025) 112003

ORIGINAL RESEARCH

User experience of Tool for Addressing Conflicts of Interest in Trials (TACIT) prototype: interview and questionnaire study

Erlend Faltinsen[a,b,c], Lasse Østengaard[a,b], Isabelle Boutron[d,e], Lesley A. Stewart[f], Asbjørn Hróbjartsson[a,b], Andreas Lundh[a,b,g,*]

[a]*Cochrane Denmark and Centre for Evidence-Based Medicine Odense, Department of Clinical Research, University of Southern Denmark, Odense, Denmark*
[b]*Open Patient data Explorative Network (OPEN), Odense University Hospital, Odense, Denmark*
[c]*Lofoten and Vesterålen District Psychiatric Centre, Nordland Hospital, Vesterålen, Norway*
[d]*Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Center for Research in Epidemiology and StatisticS (CRESS), Paris, France*
[e]*Cochrane France, Hôpital Hôtel-Dieu, Paris, France*
[f]*Centre for Reviews and Dissemination, University of York, York, UK*
[g]*Department of Respiratory Medicine and Infectious Diseases, Copenhagen University Hospital - Bispebjerg and Frederiksberg, Copenhagen, Denmark*

Accepted 24 September 2025; Published online 27 September 2025

Abstract

Objectives: To describe and analyze the user experience of a late-phase prototype of Tool for Addressing Conflicts of Interest in Trials (TACIT). The tool was developed primarily for guiding systematic reviewers in collecting, processing, and interpreting information on conflicts of interest in included randomized trials.

Study Design and Setting: A prototype tool was piloted in two settings: 1) We contacted potential participants through our network, social media, and snowball sampling. Twenty-two experienced Cochrane reviewers were invited to a virtual training module applying the tool on a preselected set of trials, and feedback was collected both during a 60-minute group session and in individual semi-structured interviews of 25-55 minutes; 2) We contacted corresponding authors of systematic review protocols from 2022. Authors of 115 protocols were invited to apply the tool on trials included in their own reviews, and feedback was collected using an online questionnaire. Recordings from the group session and interviews were transcribed and imported to Nvivo 12. Feedback was analyzed using thematic cross-case analysis. Questionnaire data were analyzed quantitatively. Based on our experience, we developed a short guide for planning of study appraisal tool piloting.

Results: Eleven experienced Cochrane reviewers and 14 systematic review teams provided feedback. Feedback was organized in four themes: general impressions, tool concepts, content and layout, and practical usage. Users generally had a favorable overall impression of the tool but suggested enhancing user friendliness and more guidance on how to make judgments based on limited conflicts of interest information and when it was reasonable to stop searching for additional information. Of 14 systematic review teams, two rated the overall impression as excellent, 11 as good, one as neutral, and no ratings of poor or very poor. The guide included nine key issues to consider when planning study appraisal tool piloting.

Conclusion: Users of a prototype tool for addressing conflicts of interest in trials generally had a positive experience, but more guidance allowing ease of use and how to deal with limited information on conflicts of interest was suggested. The feedback guided a subsequent tool adjustment, including an Excel-based application, and we provide a short guide for planning piloting of future study appraisal tools.

Keywords: Conflicts of interest; Funding; Randomized trial; Study appraisal tool; Pilot; Cochrane review; Systematic review

1. Introduction

Conflicts of interest are frequent in clinical trials [1,2] and may occur, for example, when a trial is funded by a company manufacturing one of the trial interventions, when trial researchers have financial ties to manufacturers, or when researchers have strong professional interests [3]. Conflicts of interest may impact trial results and

Funding: This research was financed through in-house resources and did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sectors.

* Corresponding author. Cochrane Denmark and Centre for Evidence-Based Medicine Odense (CEBMO), Department of Clinical Research, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark.

E-mail address: alundh@health.sdu.dk (A. Lundh).

https://doi.org/10.1016/j.jclinepi.2025.112003

Plain Language Summary

This study describes user experience of a prototype of a tool for addressing conflicts of interest in randomized clinical trials included in systematic reviews. Eleven participants who had published multiple Cochrane Reviews gave feedback in interviews after using the tool in a virtual training module. Fourteen additional review teams gave feedback via a questionnaire after having used the tool in their own systematic reviews. Users generally had a positive experience, but adaptations to make the tool easier to use were suggested. The feedback guided a subsequent tool adjustment, including development of an Excel-based application. Based on our experience, we also provide a short guide for planning piloting of future study appraisal tools.

conclusions [4−6], but there is currently limited guidance on how readers of trial publications should address conflicts of interest when interpreting trial results, for example, in the context of a systematic review [7,8].

To address this, we developed the Tool for Addressing Conflicts of Interests in Trials (TACIT) for use in Cochrane and other systematic reviews [3]. TACIT provides a framework for collecting, processing, and interpreting conflicts of interest information to make a reasoned judgment about whether there is a cause for notable concern about conflicts of interest [3,8]. In line with the framework proposed by Whiting et al. [9], we piloted a late phase version of a prototype of the tool on potential users to learn from their experiences and inform the final development of the tool.

Involvement of potential tool users has been a part of the development of other study appraisal tools [10−15]. For example, prototypes of the RoB 2 tool were piloted in two phases by systematic reviewers and then by various users after a public prototype was released [10]. Piloting of study appraisal tools has varied considerably in how early in the tool development process piloting was done, the type of users involved, the context and setting, and how feedback was collected. Piloting is usually only briefly described in the tool development sections of the core tool publications [10−15]. We therefore decided to provide a more detailed description of the methods and results from piloting a prototype of TACIT and describe potential issues to consider when planning piloting of study appraisal tools to serve as a potential blueprint for future study appraisal tool developers.

The objectives of this study were a) to describe user experience on i) the overall usefulness of the prototype tool, including the strengths and limitations of the tool structure and ii) the usefulness of the accompanying tool guidance documents; and b) to provide a practical guide for planning of study appraisal tool piloting.

2. Methods

2.1. Study design

The study is based on a preregistered study protocol (see https://osf.io/pgx7e and Supplementary Files S1 and S2). User experience was collected in two settings: i) interviews of experienced Cochrane reviewers participating in a virtual training module and ii) online questionnaire of systematic reviewers using the prototype tool in their own reviews.

2.2. Tool for addressing conflicts of interest in trials (TACIT)

TACIT requires users to collect information on trial funding and trial researchers' conflicts of interest to make a judgment of concern about conflicts of interest in a trial. Users are recommended searching for information beyond trial publications and disclosure forms, for example, trial protocols and registry information. The judgment involves three levels: trial funders, the trial's primary academic researchers (ie, nonindustry authors), and overall trial level (Fig 1). Notable concern indicates that funders (including their employees) or primary academic researchers had both: (1) important conflicts of interest and (2) important involvement in the trial. To address the problem with missing information on funding and conflicts of interest, TACIT also includes an assessment of the quality of information on which the judgment of concern was based. Assessments are supported by the use of guidance questions similar to the signaling questions used in RoB 2 and ROB-ME [10,11]. The TACIT assessment is guided by the use of a template for completing the tool assessment termed the "TACIT Grid" and a detailed guidance document (Supplementary Files S3) [16].

The prototype tool used in this study was developed by the TACIT Steering Group (AH, AL, IB, and LAS) together with the TACIT Working Group (www.tacit.one) in an iterative process and informed by the results of three TACIT supporting projects [7,17,18].

2.3. Experienced Cochrane reviewers participating in virtual training module

From November 2022 to February 2023, we recruited researchers who had authored at least three Cochrane reviews and who were either first or last author of at least one review or perceived by us as an experienced methodologist.

Recruitment included purposive sampling through direct e-mail exchange, posting in the newsletter of the Cochrane Bias Methods Group, the TACIT website (www.tacit.one), social media, and snowball sampling (eg, one participant

What is new?

Key findings

- A prototype of a Tool for Addressing Conflicts of Interests in Trials (TACIT) was piloted by experienced Cochrane reviewers and systematic reviewers of ongoing reviews.
- User generally had a positive experience, but adaptations to make the tool easier to use were suggested.

What this adds to what was known?

- This study provides a detailed summary of user feedback when piloting a study appraisal tool in different contexts and settings, which may be helpful to other study appraisal tool developers.
- A short guide for planning piloting of future study appraisal tools is presented.

What is the implication and what should change now?

- Piloting of future study appraisal tools should involve varying types of expected users applying the tool in different contexts and settings as their feedback may complement each other.
- Future studies should investigate how different characteristics, such as type of users, context, and setting, impact the usefulness of feedback in study appraisal tool piloting.
- The feedback in this study guided a subsequent adjustment of TACIT, including an Excel-based application, and an unpublished version of the tool was publicly released in September 2025.

recommending another person for participation). We aimed for at least 10 participants and tried to ensure diversity by including both researchers and clinicians and participants from different clinical fields and geographical regions.

All eligible participants who agreed to participate received a training package by e-mail (Supplementary Files S4), which was available on Nextcloud (Nextcloud GmbH). In preparation, participants watched a 30-minute training video, read training materials, and used the tool on a preselected trial [19]. On March 8, 2023, we held a 1-hour virtual group feedback session via Zoom (Zoom Communications Inc.) and moderated by one of the investigators (EF) and with an introduction by AH (AL also attended). The participants were free to give feedback at any time. The session was video recorded, and additional feedback was collected using the Zoom chat function. After the meeting, participants received two additional trials for assessment prior to interviews [20,21].

Participants then took part in individual semi-structured interviews via Zoom (also video recorded) in March 2023. The interviews were carried out by one of the investigators (EF), a psychologist working as a research assistant, with training in qualitative research methods. He had no strong beliefs about the interview topic, and the participants and the interviewer had no prior knowledge of each other.

The group session was informed by seven predefined themes, which also made the basis for an interview guide (Supplementary Files S5 and S6). The interviewer asked open-ended questions supplemented by specific prompts, made notes, and tried to allow for an open dialog by delving into comments not otherwise directly related to the predefined themes. The first interview worked as a pilot resulting in minor adjustments of the interview guide.

2.4. Systematic reviewers using prototype tool in systematic reviews

From November 2022 to March 2023, we recruited systematic review teams of ongoing systematic reviews of any type of intervention with at least five eligible randomized trials (one review only containing three trials and four observational studies were included due to recruitment challenges). We contacted corresponding authors of systematic review protocols published in 2022 in *BMJ Open*, *Cochrane Database of Systematic Reviews,* or *Systematic Reviews* by e-mail and contacted additional participants through snowball sampling.

We aimed to include participants from 15 systematic reviews (10 Cochrane reviews and five other systematic reviews) inspired by previous tool pilot studies [11,13]. For each review, a team of two systematic reviewers who agreed to participate was given the same training package as used in the virtual training module and was asked to do the TACIT assessments of trials in their review by consensus agreement, preferably the systematic reviewers involved in trial data extraction or risk of bias assessment. They were asked to document and reflect on any reasons for disagreements (eg, lack of clarity in the TACIT grid or guidance document). The participants could report the results of the TACIT assessments in their reviews if they emphasized that it was a tool prototype and could ask any queries to the principal investigator (EF) by e-mail.

After completing assessments, one of the participants received a link to an open online questionnaire, SurveyXact (Rambøll Group A/S), and was asked to complete it on behalf of both reviewers or together. The questionnaire was based on the same predefined themes as used in the group session and interviews and included six categorical questions using a five-point Likert scale and 21

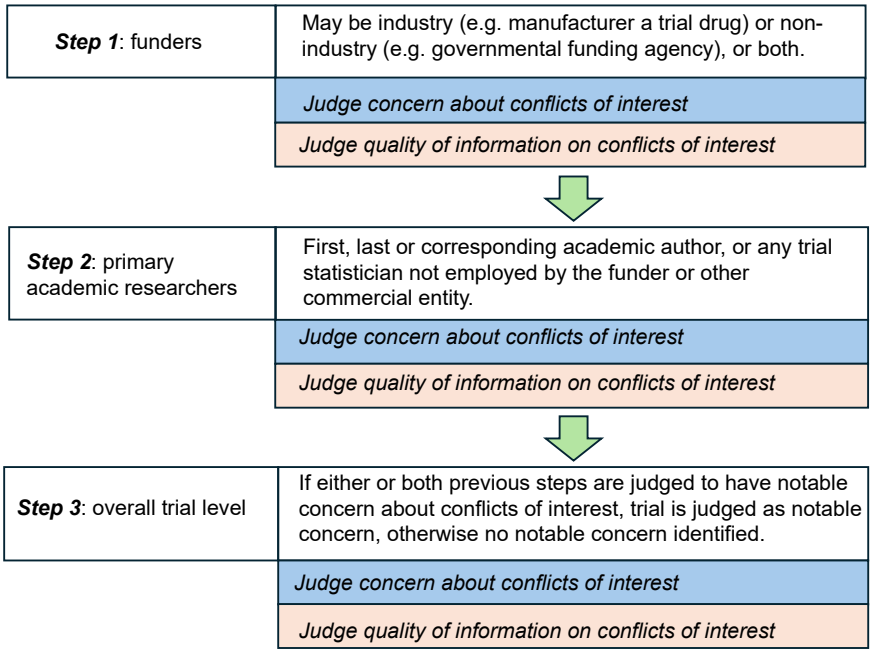

Figure 1. Outline of TACIT.

open-ended questions (Supplementary File S7). The participants were able to change their answers if needed and were required to answer all questions.

2.5. Analysis and reporting

The recorded group session and interviews were transcribed using Good Tape (Zetland Aps.), and one investigator (EF) manually corrected any transcription errors. The transcripts were not returned to participants, and the participants did not provide feedback on the data. Transcripts were imported into Nvivo 12 (Alfasoft Ltd.) and analyzed by thematic cross-case analysis using systematic text condensation [22] in four steps: i) reading transcripts to get an overview of the data; ii) identifying meaning units and coding them into groups; iii) clarifying different aspects of the coded groups (eg, creating themes and subgroups and highlighting key quotations); and iv) describing the overall user feedback on the prototype tool based on the previous steps. Two investigators (EF, AL) conducted the first and second steps, and the other steps were conducted by one investigator (EF). Another investigator (LØ), experienced in qualitative research methods, acted as a consultant during the analysis. The coding tree is included in Supplementary File S8

Data from the 21 open-ended questions in the questionnaire were quantitatively summarized, and the in-text responses were copied to a common document and analyzed qualitatively similarly to the interview data.

The qualitative part of the study was reported in line with the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist [23] (Supplementary File S9), and reporting of the questionnaire was inspired by CHERRIES (Checklist for Reporting Results of Internet E-Surveys) [24].

2.6. Guide for planning of study appraisal tool piloting

Based on developments described in other study appraisal tool publications (Supplementary File S10), our experience of piloting TACIT in this study, and through an iterative process of ongoing discussions and revisions, we developed a short guide of key issues to consider when planning piloting of a study appraisal tool.

3. Results

We contacted 22 experienced Cochrane reviewers of which 11 participated (nine in the group session and 11 in interviews) (Fig 2). We contacted corresponding authors of 115 ongoing systematic reviews of which 14 (12%) review teams participated and responded to the online questionnaire.

The characteristics of users are reported in Table 1. The users were mainly from Europe or North America (nine of 11 experience Cochrane reviewers and 15 of 27 systematic review team members). The most common characteristics for the 14 systematic reviews used in our study were Cochrane reviews ($n = 11$), related to mental health or neuroscience ($n = 6$), and included drug interventions ($n = 6$), and the reviews included a median of 19 trials (interquartile range: 16–27) (Table 2).

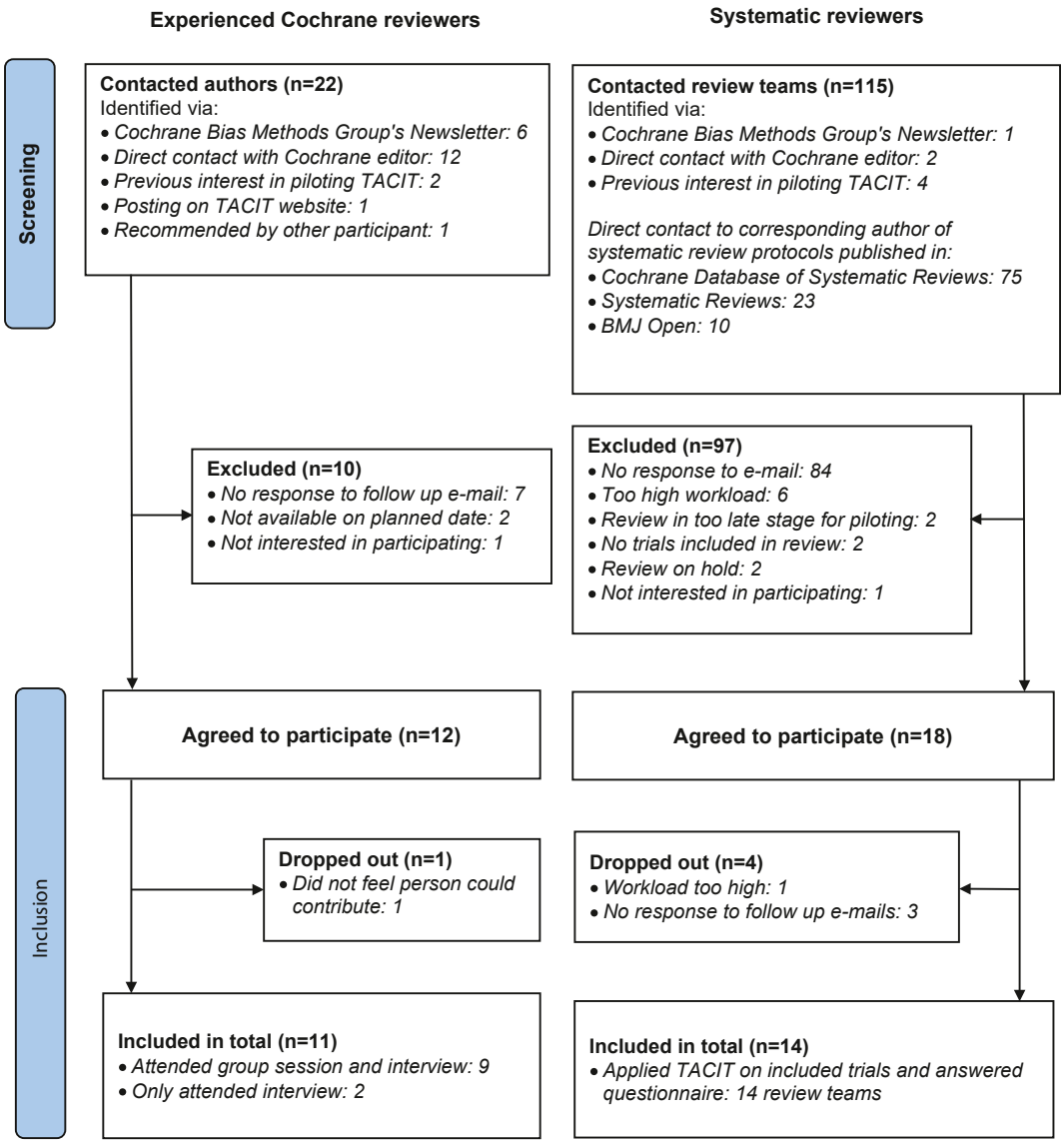

Figure 2. Flow diagram of inclusion of participants in the two stages of piloting.

3.1. Group session and interviews of experienced Cochrane reviewers participating in virtual training module

The interviews lasted a median of 29 minutes (range 25−55 minutes). Table 3 describes the feedback (with illustrative quotes) from the group session and interviews.

Based on seven predefined themes, meaning units were merged into four final themes: general impressions, tool concepts, content and layout, and practical usage (Fig 3).

3.1.1. Theme 1: general impressions

One user described a favorable impression of the tool: ''I thought it was a very well-constructed tool'' (participant 2 in group session), and another user highlighted the need to maintain balance between methodological rigor and user friendliness: ''… so there are two ways of looking at these tools that are being developed, I think. One is to have like a gold standard that people can actually learn from and another is having a tool that people can actually use in daily life'' (interviewee 2).

3.1.2. Theme 2: tool concepts

Three users addressed the challenges of dealing with missing conflicts of interest information in trial publications and redundancy of assessments. One said: ''Sometimes there's not very much information in the paper about conflicts [among academic researchers], especially with the oldest studies. So then, you'd be looking around and it's sort of like, I already know I'm going to judge this

Table 1. Characteristics of users

Category	Experienced Cochrane reviewers $N = 11$ *n*	Systematic reviewers $N = 27$[a] *n*
Primary affiliation		
Public hospital	3	8
University	5	18
Other public institution	2	1
Nonprofit organization	1	0
Continent of primary affiliation		
Africa	1	0
Asia	0	4
Europe	8	11
North America	1	4
Oceania	1	2
South America	0	6
Primary position		
Clinician	0	7
Director	1	2
Editor	1	2
Professor	2	3
Assistant professor	1	1
Researcher[b]	6	8
Grad student[c]	0	4
Editorial role in Cochrane review group		
Co-ordinating editor	4	0
Editor	4	2

Category	Median (IQR)	Median (IQR)
Cochrane review authorships	14 (6–20)	0 (0–1)
Cochrane reviews as first or last author	6 (2–10)	0 (0–0)
Other systematic review authorships[d]	13 (6–22)	1 (0 to 4)
Other systematic reviews as first or last author[d]	2 (1–4)	0 (0–2)

IQR, interquartile range.

[a] Fourteen review teams undertook TACIT assessments (13 teams with a pair of reviewers and one with a single reviewer) resulting in 27 assessors in total.

[b] Category includes mid-career researchers, like postdoctoral researchers, senior researchers, and other research associates.

[c] Category includes master's and PhD students.

[d] Data retrieved by searching for name of participant in PubMed and using the filters "Systematic Review" and "Meta-analysis" and searching for records up to end of 2022. Review protocols, Cochrane reviews, and other types of reviews or research were excluded from the count.

Table 2. Characteristics of systematic reviews

Category	$N = 14$ *n*
Type of review	
Cochrane review	11
Other systematic review	2
Part of guideline development	1
Subject area[a]	
Cancer	2
Circulation and breathing	2
Mental health and neuroscience	6
Musculoskeletal, oral, skin, and sensory	3
Public health and health systems	1
Type of intervention in review	
Behavior and education	3
Drug	6
Device	1
Nutrition and supplements	1
Screening	1
Surgery	1
Mixed[b]	1
Publication platform of review protocol	
Journal publication	13
Unpublished review protocol	1
Review included observational studies[c]	3

Category	Median (IQR)
Trials included in review	19 (16–27)

IQR, Interquartile range.

[a] Based on Cochrane's previous eight Review Networks that were closed in 2021 [https://community.cochrane.org/organizational-info/resources/resources-groups/crg-networks-portal].

[b] Two or more types of interventions (eg, drug and nondrug interventions).

[c] All reviews included at least five randomized trials except for one review, which included three trials and four observational studies.

as a notable concern because of the funder anyway … so what's the point?" (interviewee 10).

3.1.3. Theme 3: content and layout

One user commented on the amount of data collection sheets: "I was also wondering whether the appendices tables, so where you record the separate information about funding and author conflict of interest, whether those are really necessary. I thought, are they not a little bit redundant because some of the aspects actually return in the actual TACIT grid?" (participant 2 in group session). Another participant described how changing tool format and platform could improve usefulness: "I think having this as an online tool is, I think this is contemporary thing, like filling out sheets like that is not the way to do it nowadays. And it would be really good to have some kind of mouse-over … explanations on the terminology…" (interviewee 4).

3.1.4. Theme 4: practical usage

One participant described usefulness as: "I also thought it was fairly straightforward to apply." (participant 6 in

Table 3. Feedback from users in group session and interviews, ordered by themes and subgroups

Themes and subgroups	Quotes
1. General impressions	
Overall impressions	''I thought it was a very well-constructed tool'' (participant 2 in group session)
Complexity and reviewer burden	''… so there are two ways of looking at these tools that are being developed, I think. One is to have like a gold standard that people can actually learn from. And another is having a tool that people can actually use in daily life'' (interviewee 2)
	''I think it's the consumers [users of the tool]. It's the people who we are trying to impact when we are reading these papers. Actually, I think it's what lands with them. You know, have a more people-centered approach rather than a scientist-centered approach'' (interviewee 11)
2. Tool concepts	
Nonfinancial interests	''The concept of academic conflicts of interest is going to be extremely difficult, isn't it, to try to capture and to judge. But that doesn't mean it doesn't exist'' (interviewee 5)
Notable concern	''Actually, I think notable concern is more like natural language [than risk of bias]. So, I think people would understand — it's a notable concern. It's a very business term. You know, when you, when you talk about in accounting, a going concern'' (interviewee 11)
Missing information	''Instead of having sufficiency of information[a] as a separate variable, what if we just have you know unclear concerns about conflict of interest, you know if you have those two versions, how would users react to it?'' (interviewee 2)
	''Sometimes there's not very much information in the paper about conflicts [among academic researchers], especially with the oldest studies. So then, you'd be looking around and it's sort of like, I already know I'm going to judge this as a notable concern because of the funder anyway … so what's the point?'' (interviewee 10)
	''We know there's enough information out there in the world now that if it's not reported, I'm concerned. So, there you go, it's notable concern … Let's default to the most conservative efficient pathway'' (interviewee 11)
Funders and academic researchers	''It didn't make too much sense to me to concentrate on the academic authors only. Because if the senior author is industry employed, the influence is probably so strong that it overrules all other authorships'' (interviewee 4)
How assessments inform review analysis	''How do you envisage the review authors, other than doing a sensitivity analysis, really operationalizing what they find if they end up with notable concerns?'' (interviewee 5)
3. Content and layout	
Data collection sheets	''I was also wondering whether the appendices tables, so where you record the separate information about funding and author conflict of interest, whether those are really necessary. I thought, are they not a little bit redundant because some of the aspects actually return in the actual TACIT grid?'' (participant 2 in group session)
TACIT Grid	''I wondered whether the TACIT Grid could be made into a single page as a flow diagram'' (interviewee 5)
Instructional video	''I thought the video was so clear that there wasn't really any need to refer to the other guidance documents. Once I'd watched the video, I was like, oh, I understand what I'm doing'' (participant 3 in group session)
Tool format and platform	''I think having this as an online tool is, I think this is contemporary thing, like filling out sheets like ie not the way to do it nowadays. And it would be really good to have some kind of mouse-over … explanations on the terminology…'' (interviewee 4)
4. Practical usage	
Applying tool to trials	''I also thought it was fairly straightforward to apply'' (participant 6 in group session)
Making judgments and assumptions	''I suppose it just kind of highlights the inherent challenge, right, that we are going to have to make some judgments … at some point you're having to make some sort of probability assessment, right. That's just reality, I guess'' (interviewee 3)

Footnote: [x] refers to our clarification of what the interviewee meant.

[a] In the current unpublished version of TACIT, the domain ''sufficiency of information'' has been changed to ''quality of information''.

group session). Another participant described the level of subjectivity in assessments: ''I suppose it just kind of highlights the inherent challenge, right, that we are going to have to make some judgments … at some point you're having to make some sort of probability assessment, right. That's just reality, I guess'' (interviewee 3).

3.2. Questionnaire of systematic reviewers after using prototype tool in systematic reviews

For all six categorical questions, at least 10 out of 14 review teams rated the tool as either good or excellent (Fig 4). The approach for judging sufficiency of information

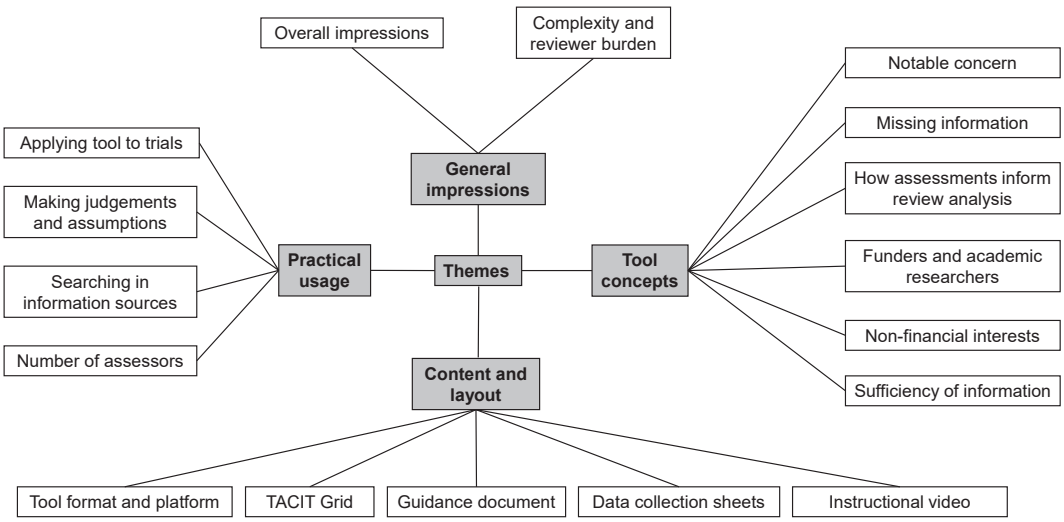

Figure 3. Thematic map of feedback from users in group session, interviews, and questionnaire.

Table 4. Feedback from users in questionnaire, ordered by themes and subgroups

Themes and subgroups	Quotes
1. General impressions	
Overall impressions	''The tool offers a level of granularity in the evaluation of potential conflicts of interest is very useful'' (review team 4)
2. Tool concepts	
Missing information	''[TACIT should] have a set of criteria for what info should reasonably be presented in a journal based on the era that it was published'' (review team 2)
Notable concern	''We develop guidelines, and our lead authors need to be free from any conflicts of interest: it is more and more difficult to find authors, unless they are really very young. I wonder if with the systematic application of the tool, we eventually will find that all the papers or the majority of the papers have a noticeable concern'' (review team 4)
Sufficiency of information[a]	''Binary classification of judging sufficiency of information was good'' (review team 1)
3. Content and layout	
TACIT Grid	''I would suggest adding a friendly flowchart or scheme to get through the evaluation (assuming users have read the well-detailed guidance)'' (review team 14)
Guidance document and instructional video	''You should consider specific cases within the guidance, such as the assessment of prepublished or no published (yet posted in a registry) results; that would be very helpful for SR authors'' (review team 14)
	''The video could have gone more in depth in what qualifies as 'lack of information'. The guidance document, however, did clarify that in a satisfactory manner'' (review team 9)
Tool format and platform	''Use an online portal to make assessments or develop an Excel spreadsheet like RoB 2'' (review team 2)
4. Practical usage	
Searching in information sources	''Although it is mentioned throughout the guidance, it is not clear enough when to stop consulting other sources of information when searching for important conflicts of interest of academic authors'' (review team 14)
Number of assessors	''Unlike RoB and GRADE [Grading of Recommendations Assessment, Development, and Evaluation] assessments, the opportunity for significant transcription error is far less. As a result I am not sure I would recommend duplicate (silo) working'' (review team 2)
	''One rater might not be enough to judge the complexity of each conflicts of interest-situation (a lot to keep in mind, could happen that information is overlooked), thus strong recommendation to use double ratings'' (review team 8)

Footnote: [x] refers to our clarification of what the interviewee meant.

[a] In the current unpublished version of TACIT, the domain ''sufficiency of information'' has been changed to ''quality of information''.

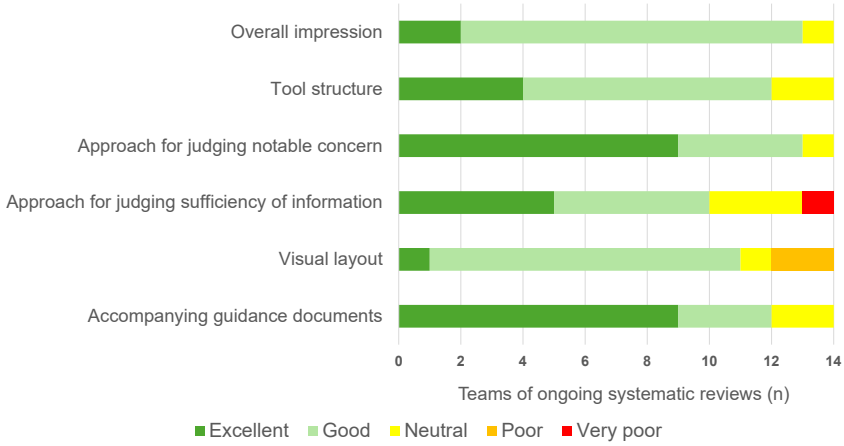

Figure 4. Ratings of TACIT in questionnaire by 14 teams of systematic reviews.

was rated very poor by one review team and the visual layout poor by two review teams (in the current unpublished version of TACIT the domain ''sufficiency of information'' has been changed to ''quality of information'').

The participants reported varying levels of time to do the assessments, from 15 to 60 minutes per trial, depending on factors like the need to retrieve additional information from other sources and that the time decreased after gaining experience with the tool.

Table 4 describes the in-text responses (with illustrative quotes) from the open-ended questions in the questionnaire. The same four final themes in the group session and interviews were used (Fig 3).

3.2.1. Theme 1: general impressions

One team described the usefulness of the tool: ''The tool offers a level of granularity in the evaluation of potential conflicts of interest that is very useful'' (review team 4).

3.2.2. Theme 2: tool concepts

One team described the challenges of missing information in older trials and whether this should be considered in the assessment: ''[TACIT should] have a set of criteria for what info should reasonably be presented in a journal based on the era that it was published'' (review team 2). The addressing of missing information by the sufficiency of information domain in TACIT was described by another team as: ''Binary classification of judging sufficiency of information was good'' (review team 1).

3.2.3. Theme 3: content and layout

One review team recommended utilizing a flow chart for the TACIT Grid: ''I would suggest adding a friendly flowchart or scheme to get through the evaluation (assuming users have read the well-detailed guidance) '' (review team 14). Another team favored a different format than a Word-based grid and datasheets: ''Use an online portal to make assessments or develop an Excel spreadsheet like RoB 2'' (review team 2).

3.2.4. Theme 4: practical usage

Two review teams disagreed as to whether the assessments should be done in duplicate or by a single assessor. One team responded: ''Unlike RoB and GRADE Grading of Recommendations Assessment, Development, and Evaluation] assessments, the opportunity for significant transcription error is far less. As a result I am not sure I would recommend duplicate (silo) working'' (review team 2). The other team responded: ''One rater might not be enough to judge the complexity of each conflict of interest-situation (a lot to keep in mind could happen that information is overlooked), thus strong recommendation to use double ratings'' (review team 8).

3.3. Guide for planning of study appraisal tool piloting

The short guide describes key issues to consider when planning study appraisal tool piloting and the pros and cons of different strategies (Table 5).

4. Discussion

Users of a prototype of TACIT generally had a favorable impression of how the tool was structured but stressed the need for improving ease of use and more guidance on how to deal with limited conflicts of interest information. They also highlighted ways to make assessments more time-efficient by simplifying the tool structure, bypassing steps, and using a more user-friendly tool platform. Based on our experience, we provide guidance for key issues to consider in future piloting of study appraisal tools.

4.1. Strengths and challenges

This study combined two methods for collecting feedback on using a prototype of TACIT. We included users

Table 5. Key issues to consider when planning study appraisal tool piloting

Issue	Examples	Pros and cons of different strategies
When to pilot	Earlier vs later in development process	Earlier piloting helps identifying problems early on, but users may be overwhelmed by the number of problems if tool is in too rough a format and then overlook important problems.
Rounds of piloting	Single or more	Several rounds of piloting can identify problems in different phases of the development process, but it is more time-consuming and may delay development.
Number of users	Few (1-5), mid (around 10), or many (20+)	The more users, the broader the representation and the higher chance of identifying key problems. However, many users are more time-consuming, and we suggest only a few users for piloting in earlier phases of the development.
Types of users	Novice systematic reviewers, experienced systematic reviewers, or methodologists	Experienced systematic reviewers and methodologists are better suited for identifying methodological issues such as inconsistencies in assessment algorithms or concepts. Novice systematic reviewers are particularly suited for identifying practical problems, such as poor guidance documents and lack of user-friendliness, and are more suited in later phases of the development.
Context	Fixed or user-tailored set of studies	Using a fixed set of studies, eg, as part of a workshop, helps to quantify how common tool problems are but may overlook problems in types of studies beyond the sample. Users piloting tool on studies included in their ''own'' systematic review ensures broader piloting, but it may be more difficult to assess if tool problems are generic or topic-related. Using a user-tailored approach is more time-consuming, has higher risk of dropouts, and may delay development.
Setting	Face to face or remote	Direct involvement with users in workshops allows for more informal feedback but requires more resources and will also limit the pool of users to the attendees at a specific event (eg conference). Online webinars or other remote piloting allows including a broader pool of users and requires fewer resources.
Timespan	1-day workshop or 3-month period	Workshops and similar short events typically only allow piloting tool on a fixed set of few studies. However, feedback can be easily collected in shorter time and requires fewer resources.
Collection of feedback	Informal oral/written feedback, questionnaire, or interviews	Structured feedback through a questionnaire or interviews ensures that predetermined topics are explored but may be a barrier for participation and be more time-consuming.
Use of public version	Public release of tool on website	Public release of tool on a website will allow feedback from a broad set of users but may be time-consuming and delay tool development.

from diverse geographical locations, professional backgrounds, methodological expertise, reviewer experience, and positions (eg, clinicians, researchers, and graduate students) and using the tool in different contexts and settings (ie, both on a preselected set of trials in the virtual training module and on trials included in different systematic reviews). This allowed a nuanced, in-depth critique to further refine the tool.

The main challenge is that in both the interviews and questionnaire, our sampling was based on a purposive selection of participants and snowballing. In the questionnaire, only 12% of the review teams initially approached participated. This selected group of users may represent people with special interests in conflicts of interest research, trial and systematic review methodology, tool development, and a certain level of experience as systematic reviewers. The feedback may therefore not represent the experience of an average user. Further, the tool was only used on a small number of trials, which limits the assessment of tool performance beyond this sample. However, it was not the aim of our study to determine the frequency of events or issues, and our qualitative approach allowed for detailed feedback which importantly informed refinement of the tool. Another challenge in our approach was that the qualitative analysis required subjective assessments and selection of which feedback best reflected the data and which themes were present. Our own involvement in the development of TACIT may have influenced this process, and to minimize such partiality, the principal investigator and main person involved in the analysis have not been involved in the development of the actual tool.

4.2. Other studies

To our knowledge, this is the first paper that describes in detail the conduct and user experience of piloting a study appraisal tool. The development of other study appraisal tools, such as AMSTAR 2, PROBAST, RoB 2, ROBINS-I, ROB-ME, and QUADAS-2 [10–15], also involved users with varying levels of methodological and reviewer experience and in different contexts and settings, similar to our study. For example, the development of ROB-ME included

feedback from both webinar participants and systematic reviewers using the tool in reviews [11]. However, the development of other tools has mainly relied on written feedback in the form of questionnaires or comments or verbal feedback delivered at workshops or webinars. Our study supplemented this with feedback from individual semi-structured interviews, and while the feedback from interviews related to the same themes as feedback in our questionnaire, it was more detailed and specific to tool concepts. Furthermore, previous reports on other tools have focused mainly on the structure and content of tools and only briefly described the actual tool development process, providing little detail on how piloting was done, how feedback was collected, and the issues raised in the feedback [10−15]. This report therefore adds important knowledge to the field of study appraisal tool development.

4.3. Implications

This study stresses the importance of involving a range of expected users in tool development as their feedback may complement each other. The use of individual semi-structured interviews allows for detailed feedback that is useful in tool revision. Further, this study provides guidance and outlines some of the key issues to consider when piloting study appraisal tools and may be used by tool developers when planning tool development.

While users have generally been involved in the development of recent study appraisal tools, there seems to be little empirical knowledge on how best to involve them in the process. Future studies should investigate how different characteristics such as type of users, context, setting, and how feedback is collected, impact on the usefulness of feedback, for example, by randomizing expected tool users to different contexts and methods for collecting feedback, similar to the approach used in the study of peer review [25] and reporting guidelines [26,27].

The results of this study together with three other supporting projects form the basis for the development of TACIT [7,17,18]. Based on the feedback, the tool has been simplified, includes more guidance for how to deal with missing information, allows the possibility of bypassing certain tool steps, and an Excel-based application of the tool has been developed. The collected feedback and subsequent tool revision informed a virtual consensus meeting by the TACIT Working Group in May 2025. Based on the discussions at the consensus meeting, the tool was revised and an unpublished version was released in September 2025 for a phase of public feedback [16].

5. Conclusion

Feedback from users of a prototype of a tool for addressing conflicts of interest in trials was generally positive, both in interviews of experienced Cochrane reviewers and in a questionnaire of systematic review teams, but more guidance allowing ease of use and how to deal with limited information on conflicts of interest was suggested. The feedback guided a subsequent tool adjustment, including an Excel-based application. Our study highlights how involvement of various users, using the tool in different setting and collecting feedback through different methods, may complement each other and contribute to tool development. Further, based on our experience, we have developed a short guidance for key issues to consider in piloting of future study appraisal tools.

Ethics and data protection

An agreement on safe processing of personal data of the study participants was granted by the Research and Innovation Organization (RIO) at University of Southern Denmark prior to conducting the study. All personal data were handled with confidentiality in accordance with Danish law and the data were either de-identified or stored safely, and all data considered personally identifiable information will be deleted before January 1, 2026. RIO confirmed that according to Danish law, questionnaire and interview studies, not using human biological materials, do not require ethics approval and the study was therefore exempt from other approval procedures (eg, an institutional review board).

The lead author (EF) confirms that the manuscript is an honest, accurate, and transparent account of the study being reported. No important aspects of the study have been omitted and any discrepancies from the study as originally planned have been reported and described.

CRediT authorship contribution statement

Erlend Faltinsen: Writing − review & editing, Writing − original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Lasse Østengaard:** Writing − review & editing, Formal analysis. **Isabelle Boutron:** Writing − review & editing, Methodology. **Lesley A. Stewart:** Writing − review & editing, Methodology. **Asbjørn Hróbjartsson:** Writing − review & editing, Methodology, Conceptualization. **Andreas Lundh:** Writing − review & editing, Writing − original draft, Supervision, Methodology, Formal analysis, Conceptualization.

Declaration of competing interest

AH, AL, EF, IB, and LØ are all affiliated with national Cochrane centers and AH and IB are convenors of the Cochrane Bias Methods group. AH, AL, IB, and LAS make up the TACIT Steering Group. AH is the Assistant Editor of the Cochrane Methodology Review Group. LAS is co-convenor of

the Cochrane Individual Participant Data Meta-analysis Group. The authors declare no other relevant interests.

Acknowledgments

We thank Sarah Lensen, Muhammad Imran Omar, Tanya Horsley, Susan Brunskill, Silvia Minozzi, Bada Yang, Jack Wilkinson, Simon Lewin, Harald Herkner, Madelon van Wely, Nandi Siegfried, Fulvia Baldassarre, Norma Varela, Thomas Munder, Christiane Steinert, Kotomi Sakai, Kalyan Vemulapalli, Jisha Sarah John, Sherly Nelapati, Christopher Carswell, Ville Ponkilainen, Sheila Cyril, Nurul Husna Salahudddin, Arturo Martí-Carvajal, Mario Gemmato, Rebeca Mangabeira, Mateus Alves Borges Christino, Teemu Karjalainen, Niyati Vyas, Dria Bennett, Robert Pap, Nicolás Meza Concha, Vincente Pablo Núnez, and Sidsel Christine Buskbjerg Døssing for giving their feedback on the tool. We also thank all the other participants who took part in the study but did not respond to our request to be named in the acknowledgments.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2025.112003.

Data availability

De-identified interview and questionnaire data are available in PDF format at https://osf.io/pgx7e/files.

References

[1] Hakoum MB, Jouni N, Abou-Jaoude EA, Hasbani DJ, Abou-Jaoude EA, Lopes LC, et al. Characteristics of funding of clinical trials: cross-sectional survey and proposed guidance. BMJ Open 2017; 7:e015997. https://doi.org/10.1136/bmjopen-2017-015997.

[2] Hakoum MB, Jouni N, Abou-Jaoude EA, Hasbani DJ, Abou-Jaoude EA, Lopes LC, et al. Authors of clinical trials reported individual and financial conflicts of interest more frequently than institutional and nonfinancial ones: a methodological survey. J Clin Epidemiol 2017;87:78−86. https://doi.org/10.1016/j.jclinepi.2017.04.002.

[3] Lundh A, Boutron I, Stewart L, Hróbjartsson A. What to do with a clinical trial with conflicts of interest. BMJ Evid Based Med 2020; 25:157−8. https://doi.org/10.1136/bmjebm-2019-111230.

[4] Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L. Industry sponsorship and research outcome. Cochrane Database Syst Rev 2017; 2:MR000033. https://doi.org/10.1002/14651858.MR000033.pub3.

[5] Ahn R, Woodbridge A, Abraham A, Saba S, Korenstein D, Madden E, et al. Financial ties of principal investigators and randomized controlled trial outcomes: cross sectional study. BMJ 2017;356: i6770. https://doi.org/10.1136/bmj.i6770.

[6] Nejstgaard CH, Laursen DRT, Lundh A, Hróbjartsson A. Commercial funding and estimated intervention effects in randomized clinical trials: systematic review of meta-epidemiological studies. Res Synth Methods 2023;14:144−55. https://doi.org/10.1002/jrsm.1611.

[7] Lundh A, Rasmussen K, Østengaard L, Boutron I, Stewart LA, Hróbjartsson A. Systematic review finds that appraisal tools for medical research studies address conflicts of interest superficially. J Clin Epidemiol 2020;120:104−15. https://doi.org/10.1016/j.jclinepi.2019.12.005.

[8] Boutron I, Page MJ, Higgins JPT, Altman DG, Lundh A, Hróbjartsson A. Chapter 7: considering bias and conflicts of interest among the included studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Intervention. Version 6.5. Cochrane; 2024. Available at: https://www.cochrane.org/handbook. Accessed September 3, 2025.

[9] Whiting P, Wolff R, Mallett S, Simera I, Savović J. A proposed framework for developing quality assessment tools. Syst Rev 2017; 6:204. https://doi.org/10.1186/s13643-017-0604-6.

[10] Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366:l4898. https://doi.org/10.1136/bmj.l4898.

[11] Page MJ, Sterne JAC, Boutron I, Hróbjartsson A, Kirkham JJ, Li T, et al. ROB-ME: a tool for assessing risk of bias due to missing evidence in systematic reviews with meta-analysis. BMJ 2023;383: e076754. https://doi.org/10.1136/bmj-2023-076754.

[12] Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355:i4919. https://doi.org/10.1136/bmj.i4919.

[13] Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017;358:j4008. https://doi.org/10.1136/bmj.j4008.

[14] Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529−36. https://doi.org/10.7326/0003-4819-155-8-201110180-00009.

[15] Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med 2019; 170:51−8. https://doi.org/10.7326/M18-1376.

[16] Lundh A, Matheson A, Webster A, Chan AW, Thombs B, Akl EA, et al. Tool for addressing conflicts of interest in trials (TACIT) - guidance document. Version 2. 2025. Available at: https://osf.io/5fzv3. Accessed September 3, 2025.

[17] Østengaard L, Lundh A, Tjørnhøj-Thomsen T, Abdi S, Gelle MHA, Stewart LA, et al. Influence and management of conflicts of interest in randomised clinical trials: qualitative interview study. BMJ 2020; 371:m3764. https://doi.org/10.1136/bmj.m3764.

[18] Faltinsen E, Todorovac A, Boutron I, Stewart LA, Hróbjartsson A, Lundh A. A structured approach to information retrieval improved identification of funding and researchers' conflicts of interest in trials included in cochrane reviews. J Clin Epidemiol 2023;161:104−15. https://doi.org/10.1016/j.jclinepi.2023.06.020.

[19] Lipton RB, Croop R, Stock DA, Madonia J, Forshaw M, Lovegren M, et al. Safety, tolerability, and efficacy of zavegepant 10 mg nasal spray for the acute treatment of migraine in the USA: a phase 3, double-blind, randomised, placebo-controlled multicentre trial. Lancet Neurol 2023;22:209−17. https://doi.org/10.1016/S1474-4422(22)00517-8.

[20] A prospective, randomized, double-masked trial on radiation therapy for neovascular age-related macular degeneration (RAD study). Radiation therapy for Age-related macular degeneration. Ophthalmology 1999;106:2239−47. https://doi.org/10.1016/S0161-6420(99)90522-5.

[21] Caraco Y, Crofoot GE, Moncada PA, Galustyan AN, Musungaie DB, Payne B, et al. Phase 2/3 trial of molnupiravir for treatment of Covid-19 in nonhospitalized adults. NEJM Evid 2022;1:EVIDoa2100043. https://doi.org/10.1056/EVIDoa2100043.

[22] Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scand J Public Health 2012;40:795−805. https://doi.org/10.1177/1403494812465030.

[23] Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19:349−57. https://doi.org/10.1093/intqhc/mzm042.

[24] Eysenbach G. Improving the quality of web surveys: the checklist for reporting results of internet E-Surveys (CHERRIES). J Med Internet Res 2004;6:e34. https://doi.org/10.2196/jmir.6.3.e34.

[25] Speich B, Mann E, Schönenberger CM, Mellor K, Griessbach AN, Dhiman P, et al. Reminding peer reviewers of reporting guideline items to improve completeness in published articles: primary results of 2 randomized trials. JAMA Netw Open 2023;6:e2317651. https://doi.org/10.1001/jamanetworkopen.2023.17651.

[26] 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;14:199. https://doi.org/10.1186/s12916-016-0736-x.

[27] Yavchitz A, Ravaud P, Hopewell S, Baron G, Boutron I. Impact of adding a limitations section to abstracts of systematic reviews on readers' interpretation: a randomized controlled trial. BMC Med Res Methodol 2014;14:123. https://doi.org/10.1186/1471-2288-14-123.