
Review

Evidence appraisal: a scoping review, conceptual framework, and research agenda

Andrew Goldstein,¹ Eric Venker,² and Chunhua Weng¹

¹Department of Biomedical Informatics, Columbia University, New York, NY, USA and ²Department of Medicine, Columbia University, New York, NY, USA

Corresponding Author: Andrew Goldstein, 622 West 168th Street PH20, New York, NY 10032, USA. E-mail: ag3304@cumc.columbia.edu. Phone: +1 215-805-0082

Received 23 December 2016; Revised 25 March 2017; Accepted 18 April 2017

ABSTRACT

Objective: Critical appraisal of clinical evidence promises to help prevent, detect, and address flaws related to study importance, ethics, validity, applicability, and reporting. These research issues are of growing concern. The purpose of this scoping review is to survey the current literature on evidence appraisal to develop a conceptual framework and an informatics research agenda.

Methods: We conducted an iterative literature search of Medline for discussion or research on the critical appraisal of clinical evidence. After title and abstract review, 121 articles were included in the analysis. We performed qualitative thematic analysis to describe the evidence appraisal architecture and its issues and opportunities. From this analysis, we derived a conceptual framework and an informatics research agenda.

Results: We identified 68 themes in 10 categories. This analysis revealed that the practice of evidence appraisal is quite common but is rarely subjected to documentation, organization, validation, integration, or uptake. This is related to underdeveloped tools, scant incentives, and insufficient acquisition of appraisal data and transformation of the data into usable knowledge.

Discussion: The gaps in acquiring appraisal data, transforming the data into actionable information and knowledge, and ensuring its dissemination and adoption can be addressed with proven informatics approaches.

Conclusions: Evidence appraisal faces several challenges, but implementing an informatics research agenda would likely help realize the potential of evidence appraisal for improving the rigor and value of clinical evidence.

Key words: critical appraisal, post-publication peer review, journal clubs, journal comments, clinical research informatics

BACKGROUND AND SIGNIFICANCE

Clinical research yields knowledge with immense downstream health benefits,^{1–4} but individual studies may have significant flaws in their planning, conduct, analysis, or reporting, resulting in ethical violations, wasted scientific resources, and dissemination of misinformation, with subsequent health harm.^{1,5–8} While much of the evidence base may be of high quality, individual studies may have issues, including nonalignment with knowledge gaps or health needs, neglect of prior research,^{1,7} lack of biological plausibility,⁹ miscalculation of statistical power,^{10,11} poor randomization, low-quality blinding,¹²

flawed selection of study participants or flawed outcome measures,^{7,13–15} ethical misconduct during the trial (including issues with informed consent,^{16,17} data fabrication, or data falsification),^{18,19} flaws in statistical analysis,²⁰ nonreporting of research protocol changes,⁷ outcome switching,²¹ selective reporting of positive results,⁵ plagiarized reporting,^{19,22} misinterpretation of data, unjustified conclusions,^{23,24} and publication biases toward statistical significance and newsworthiness.⁷

In order to prevent, detect, and address these issues, several mechanisms exist or have been promoted, including reforming incen-

tives for research funding, publication, and career promotion^{1,5-8,25}; pre-registering trial protocols^{5,26-28}; sharing patient-level trial data^{5,23,29-31}; changing journal-based pre-publication peer review^{32,33}; performing meta-research^{34,35}; performing risk of bias assessment as part of meta-analysis, systematic review, and development of guidelines and policies^{23,36-38}; and expanding reproducibility work via data reanalysis³⁹ and replication studies.⁴⁰ However, despite these measures, the published literature often lacks value, rigor, and appropriate interpretation. This is occurring during a period of growth in the published literature, leading to information overload.⁴¹⁻⁴³ Further, there is a large volume of nontraditional and emerging sources of evidence: results, analyses, and conclusions outside of the scientific peer-reviewed literature, including via trial registries and data repositories,^{26,27,44,45} observational datasets,^{46,47} publication without journal-based peer review,⁴⁸⁻⁵⁰ and scientific blogging.⁵¹⁻⁵³

Here, therefore, we discuss evidence appraisal. This refers to the critical appraisal of published clinical studies via evaluation and interpretation by informed stakeholders, and is also called trial evaluation, critical appraisal, and post-publication peer review (PPPR). Evidence appraisal has generated considerable interest,^{23,48,54-64} and a significant volume of appraisal is already occurring with academic journal clubs, published journal comments (eg, PubMed has more than 55 000 indexed clinical trials that have 1 or more journal comments), PPPR platforms, social media commentary, and trial risk of bias assessment as part of evidence synthesis.

While evidence appraisal is prevalent, downstream use of its generated knowledge is underdefined and underrealized. We believe evidence appraisal addresses many important needs of the research enterprise. It is key to bridging the T2 phase (clinical research or translation to patients) and T3 phase (implementation or translation to practice) of translational research by ensuring appropriate dissemination and uptake (including implementation and future research planning).⁶⁵ Further, appraisal of evidence by patients, practitioners, policy-makers, and researchers enables informed stakeholder feedback and consensus-making as part of both learning health system⁶⁶⁻⁶⁸ and patient-centered outcomes research⁶⁹⁻⁷¹ paradigms, and can help lead to research that better meets health needs. As appraisal of evidence can help determine research quality and appropriate interpretation, this process is also aligned with recent calls to improve the rigor, reproducibility, and transparency of published science.⁷²⁻⁷⁴ Evidence appraisal knowledge can be envisioned, therefore, to enable a closed feedback loop that clarifies primary research and enables better interpretation of the evidence base, detection of research flaws, application of evidence in practice and policy settings, and alignment of future research with health needs.

Despite these benefits, the field of evidence appraisal is undefined and understudied. Though there is some siloed literature on journal commentary, journal clubs, and online PPPR (which we review and describe here), the broader topic of evidence appraisal, especially the uptake of knowledge emerging from this process, is unexplored and we find no reviews, conceptual frameworks, or research agendas for this important field.

Objective

In this context, our goals are to address these knowledge gaps by better characterizing evidence appraisal and proposing a systematic framework for it, as well as to raise awareness of this important but neglected step in the evidence lifecycle and component of the research system. We also seek to identify opportunities to increase the scale, rigor, and value of appraisal. Specifically, informatics is enabling transformative improvement of information and knowledge

management involving the application of large, open datasets, social computing platforms, and automation tools, which could similarly be leveraged to enhance evidence appraisal.

Therefore, we performed a scoping review of post-publication evidence appraisal within the biomedical literature and describe its process architecture, appraisers, use cases, issues, and informatics opportunities. Our ultimate goal was to develop a conceptual framework and propose an informatics research agenda by identifying high-value opportunities to advance research in systematic evidence appraisal and improve appropriate evidence utilization throughout the translational research lifecycle.

MATERIALS AND METHODS

We conducted a scoping review, including a thematic analysis. Our methodology was adapted from Arksey and O'Malley, with changes allowing for broader inclusion, iterative analysis, and more efficient implementation.⁷⁵ This method involves 3 steps: (1) identifying potential studies, (2) screening studies for inclusion, and (3) conducting thematic analysis, with collation and reporting of the findings.

Article search

To identify potential studies in the biomedical literature, we searched for any citation in Medline published by the time of our search, December 6, 2015. Our search prioritized precision while maintaining temporal comprehensiveness at the cost of recall via more comprehensive search terms. This search strategy placed emphasis on specific search terms that were more likely to identify articles relevant to evidence appraisal (Supplementary Appendix 1A). Search terms related to appraisal included PPPR, journal club, and the evaluation, assessment, or appraisal of studies, trials, or evidence. We used this narrower set of search terms in 1 citation database (Medline). Our initial search yielded 2187 citations. During screening and thematic analysis, further search terms were identified (Appendix 1B), yielding another 237 citations. Of these 2424 citations, there were 2422 unique articles.

Article screening

One researcher (AG) manually screened the titles and abstracts of all retrieved articles. Articles were included if they contained discussion of the post-publication appraisal processes, challenges, or opportunities for evidence that was either in science or biomedicine generally or specific to health interventions. Articles were excluded if they did not meet those criteria or were instances of evidence appraisal (eg, the publication of proceedings from a journal club) or discussed evidence appraisal as part of pre-publication review, meta-research, meta-analysis, or systematic review. Citations were included regardless of article type (eg, studies, reviews, announcements, and commentary). Non-English articles were excluded. Ultimately, this led to a total of 121 articles meeting the criteria for thematic analysis. The search and screening Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram⁷⁶ is shown in Figure 1.

Thematic analysis

Two reviewers (AG and EV) extracted themes manually from the full-text article or, if this was unavailable, from the abstract. The thematic analysis involved iterative development of the coding scheme and iterative addition of search terms for identifying new articles. The coding scheme was updated based on discussions and consensus of both reviewers, followed by reassessment of previously coded articles for the new codes. Coding was validated with discrepancies adjudicated by discussion until consistency and consensus were

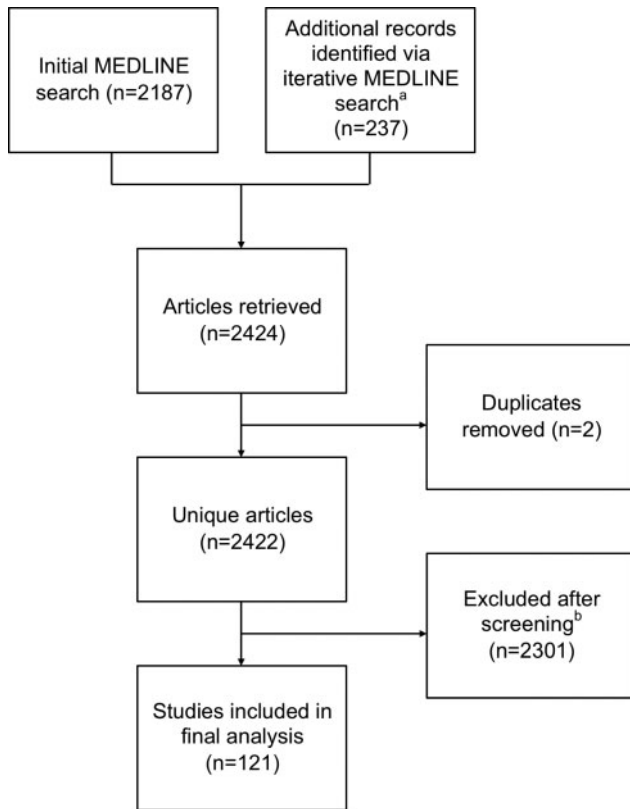


Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram demonstrating article screening process. (A) During article screening and initial thematic analysis, additional search terms were identified and a further search was performed. (B) Screening was performed on the title and, as available, the abstract of retrieved articles, with full text reviewed if available and further ambiguity remained.

achieved. Final themes were then grouped into categories. These themes and categories were then used to develop a de novo conceptual framework by author consensus.

RESULTS

Descriptive statistics

The 121 articles included in this review were published in 90 unique journals, with each journal on average having 1.34 articles included in the study (standard deviation 1.00). The maximum number of articles from a single journal (*Frontiers in Computational Neuroscience*) was 9. There were 4 articles from the *Journal of Evaluation in Clinical Practice*, and 3 each from *BMC Medical Education*, *Clinical Orthopedics and Related Research*, the *Journal of Medical Internet Research*, and *Medical Teacher*. The number of articles published by year is found in Figure 2. We were able to access the full text for 119 articles and used only the abstract for the remaining 2.

The articles spanned a range of publication types: 24 reviews, 18 trials, 3 observational studies, 21 surveys, 44 editorials and perspectives, 7 comments and letters to the editor, and 4 journal announcements (Figure 3).

Sixty-eight distinct themes forming 10 categories (A–J) were identified from these 121 articles. An average of 7.16 themes were found in each article (standard deviation 3.44). Themes occurred 866 times in total, with a mean of 12.55 instances per theme (standard deviation 15.71). The themes, theme categories, and articles in which themes were identified are provided in Table 1 and described below.

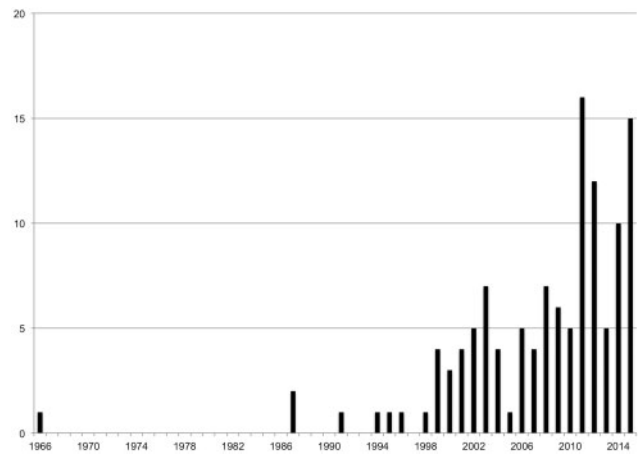


Figure 2. Number of articles by publication year.

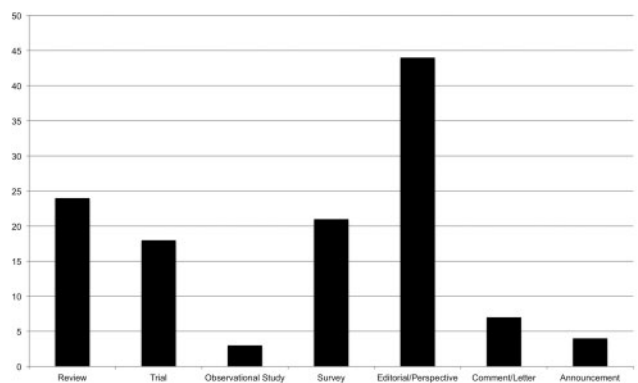


Figure 3. Number of articles by publication type.

Category A: Appraisal initiation promoters

Many articles discussed why appraisal was or was not initiated (category A). Appraisal is often required for student coursework (theme 1), required or encouraged for continuing education (theme 2), or performed to answer evidentiary questions at the point of practice (theme 3), including clinical practice, program development, and policymaking. These themes were frequently discussed related to a broader theme of evidence-based practice. Most of these instances were appraiser-centric in that they were focused on the appraisal process benefiting the practitioner's ability to perform daily tasks via a better personal critical understanding of the evidence.

In contrast to appraiser-centric appraisal as part of improving appropriate evidence application, a significant share of appraisal was motivated by the goal of improving the rigor of scientific evidence and its interpretation (theme 4) and was more commonly focused on the research system. These articles highlighted fields of open science, such as open-access publishing and open evaluation. As some open-access publications do not undergo pre-publication peer review, post-publication peer review was described as the only review process for these articles (theme 5).

When appraisal happens, certain attributes of the process are perceived to motivate increases in the scale or quality of appraisal. Some meetings, such as journal clubs, are scheduled regularly, which serves to both enhance attendance and develop institutional culture and habits of evidence appraisal (theme 6). Meetings being mandatory (theme 7) was discussed in conflicting ways. By some, mandatory sessions were considered to increase attendance, and thus the scale and richness

Table 1. Thematic analysis

Themes	References	
1 (A, D)	Coursework requirement	77–90
2 (A, D)	Motivated by continuing education	56,60,81,83,85,91–131
3 (A, D)	Point-of-practice evidence need	85,98,116,118,132,133
4 (A)	Motivation to improve science	48,50,54,56–58,86,134–140
5 (A)	Post-publication peer review only	48–50,54
6 (A)	Meeting regularity	77–79,81–83,85,87,89,91,92,94–107,109,110,113,115,116,119–122,125,126,129,131,140–165
7 (A)	Meeting voluntariness	77,78,80,83,87,91,95–97,102,103,105,106,109,114,115,120,140–142,145,146,151–153,157,161,164,166,167
8 (A)	Status incentives	48–50,56,127,137,138,168
9 (A)	Other incentives	50,56,84,89,91,92,94–99,102,103,105–107,109,114,115,120,127,129,130,141,142,145,147,151,153,155,161,164,166,167
10 (B)	Time burden	50,56,83,85,90,91,93,94,98–100,105,107,110,112,115,116,119,122,127–129,132,139,151,154,156,169
11 (B)	Use of existing evidence appraisals	85,93,115,127,143,170
12 (B)	Journal access barriers	60
13 (B)	Scarce funding	138,139
14 (C)	Landmark studies	85,94,95,103,109,147,156,158
15 (C)	Recent studies	62,86,89,91,94–99,102,103,105,107,108,113,116,119,120,122,124,128,142,144,147,151,152,158,163,166,167
16 (C)	Practice questions	77,79,81,83,85,87,88,91,92,96–98,103,105,107,110,112,114–116,118,120–123,126,127,131–133,143,144,147,148,150,155,156,161,166,169–173
17 (C)	Recommendation systems	50
18 (D)	Lay stakeholders	60,168
19 (D)	Multidisciplinary team	89,93,107,111,131,152,161,164,168
20 (D)	Appraiser identification	48–50,56,57,59,60,134,137,138,168,174
21 (E)	Guideline development tools	131,132,135,144,170,175
22 (E)	Cochrane tools	118,131,170,173,175
23 (E)	Delphi list tool	173
24 (E)	EQUATOR tools	82,109,118,154,171,176
25 (E)	<i>JAMA's Users' Guide to the Medical Literature</i>	87,97,118,120,122,132,143,144,175
26 (E)	Tools from the Centre for Evidence-Based Medicine, Oxford	83,87,92,97,98,110,118,120,122,127,132,143,155,172,177
27 (E)	Critical Appraisal Skills Programme (Oxford) tools	118,122
28 (E)	Unspecified structured tools	77,79,81,82,85,88–97,101,103,104,107,117,120,121,129,132,133,136,141,142,144,145,147,148,153,155,161,171,173,178–180
29 (E)	Unspecified journal club format	78,80,84,85,89,91,95,97,99–106,108,111–117,119,120,123–126,128,140,142,144–153,157–160,162,164–167,178,181
30 (E)	Overall ratings	48–50,58,138
31 (E)	Journal prestige as proxy	48,50,56,168
32 (E)	Automated fraud detection	57,60
33 (F,G)	In-person group	56,77,78,80–85,87–92,95–117,119–126,128,129,131,133,140–162,164–168,172,179–181
34 (F, G)	Social media	54,59,99,112,130,131,137,163,174,177,182–185
35 (F, G)	PPPR platform	48,50,54,56–60,137,138,168,174,182,186
36 (F, G)	Blog	48,49,54,59,60,99,131,137,164,168,174,182,186
37 (F, G)	Journal site message board	49,50,54,56,58,60,62,86,137,174,186
38 (F, G)	Private, virtual journal club	80,93,96,97,99,100,103,114,120,122,131,141,142,155,156,164,165,180
39 (F)	Journal comment	61,86,89,127,137,138,174,182,186
40 (F)	Individually, unspecified	56,60,79,86,94,118,121,127,133,134,139,150,153,168,170,171,173,178
41 (F)	Institutional professional appraisal service	169
42 (G)	Facilitation	89,91–94,96,98–100,102–104,109,111,114,116,123,129,131,138,141,145,155,174,177,179,182
43 (G)	Debate format	156,179
44 (G)	Generative-comparative format	97,155
45 (G)	Indefinite appraisal	48,49,93,134,137,138,168
46 (G)	Collaborative editing format	60,99,137
47 (H)	Negativity and positivity biases	59,134,139
48 (H)	Incivility	86,134,138,168,174,182
49 (H)	Enabling low quality	49,54,61,168,174
50 (H)	Platform underuse	49,50,54,56,138,174
51 (H, D)	Inter-appraiser reliability and individual appraiser performance	48,50,91,136,168
52 (H, E)	Appraisal dimensions gap	135,136,175,176
53 (H, E)	Heterogeneous needs by field	132,136,175

(continued)

Table 1. continued

Themes	References	
54 (I)	Unrecorded	54,56,60,110,166,168
55 (I)	Local publication or storage	77,79,83,93,97,100,103,105,110,114,120,122,147,152,155,166,172
56 (I)	Immediately published online	48,50,54,59,61,62,112,137,138,168,174,177,186
57 (I)	Journal-based publication	48,61,62,80,86,89,110,127,138,139,153
58 (I)	Indexing or globally aggregating	56,61,130,174,177,184
59 (I)	Data standards	168
60 (I)	Linking to appraised evidence	61
61 (I)	Evaluation of appraisal	48,50,57,59,62,86,138,168,177,186
62 (I)	Messaging appraisal to evidence authors	58,59,62,86,130,131,163,168,177
63 (I)	Correcting or retracting evidence	57,59,60,138,174,186
64 (I)	Focused aggregation	48,50,56,130,168,174,177
65 (J)	Worsening of information overload	54,60,168
66 (J)	Need to manage redundancy	60
67 (J)	Low volume or quality of research on evidence appraisal	85,95,101,132,139,145,181
68 (J)	Low uptake of evidence appraisal	86

Themes are listed with theme codes in parentheses; the first letter is the major theme category it is a part of, followed by additional categories it is a member of. For example, theme 1 is primarily in category A but also in category D. Theme category key: A = Initiation Promoters; B = Initiation Inhibitors; C = Article Selection Sources; D = Appraiser Characteristics; E = Tools and Methods; F = Appraisal Venue; G = Format Characteristics; H = Appraisal Process Issues; I = Post-Appraisal Publication, Organization, and Use; J = Post-Appraisal Issues and Concerns. EQUATOR = Enhancing the Quality and Transparency of Health Research; JAMA = *Journal of the American Medical Association*

of group discussion, but one article described voluntary attendance as preferred, as participants are more self-motivated to attend and more likely to participate. Social status incentives for appraising or for appraising with high quality, as well as research career incentives via academic rewards for appraisals, were described as potentially increasing participation (theme 8). Various other incentives that increase participation were also mentioned (theme 9), including offering continuing education credits, providing food, providing time for socializing, meeting local institutional requirements for work or training, and offering unspecified incentives.

Category B: Appraisal initiation inhibitors

Articles also described several inhibitors of appraisal initiation (category B). The most commonly cited inhibitor was the time burden to perform evidence appraisal and the time requirements for other work (theme 10). One result of this is that when appraised evidence is demanded, some practitioners will utilize or rely on previous appraisals made by others rather than producing their own (theme 11). Other inhibitors of appraisal include lack of access to full-text articles (theme 12) and lack of explicit funding support for performing appraisal (theme 13).

Category C: Article selection sources

There are many rationales for or contexts in which an article might be appraised (category C). Typically, selection is task-dependent related to learning, practice, or specific research needs. For initial learning in a field, landmark studies (theme 14), also called classic, seminal, or practice-changing studies, were selected for their historical value in helping learners understand current practices and field-specific trends in research. For continuing education, articles were selected for appraisal based on recently published articles (theme 15) or from recent clinical questions (theme 16). A novel approach to article selection for continuing education was discussed: collaborative filtering or recommendation systems (theme 17), which provide a feed of content to users based on their interests or on prior evidence reviewed. For skills development-oriented appraisal,

selection of studies also aligned with classic, recent, or practice question-driven study selection. Of note, very frequently no rationale for article selection was explicitly stated.

Category D: Common and innovative appraiser characteristics

Appraisers had varied characteristics (category D). It is implicit in many themes that appraisal stakeholders include researchers and practitioners, at the student (theme 1) as well as training and post-training levels (themes 2 and 3). The stakeholders involved in a given appraisal setting were typically a relatively homogenous group of practitioners at a similar stage of training and in silos by field. However, other appraisers or groups with innovative characteristics have also been engaged and warranted thematic analysis.

Rarely, appraisal included or was suggested to include laypeople, such as patients, study participants, or the general public (theme 18). Some groups performing appraisal were multidisciplinary and included different types of health professionals or methodologists, including librarians and those with expertise in epidemiology or biostatistics (theme 19).

For post-publication peer-review platforms, disclosure of the identity of appraisers was a common yet controversial topic. There were rationales advocating for and against, and examples of anonymous, pseudonymous, and named authorship of appraisals (theme 20). The motivating factor for this was optimizing who decides to initiate, what and how they critique, how they express and publish it, and how the appraisal is received by readers.

Regarding appraiser attributes, only one significant issue was raised: individual appraiser quality and subsequent inter-appraiser unreliability (theme 51).

Category E: Appraisal tools and methods

There are many tools for appraisal (category E), mostly to support humans performing appraisal. Most evidence appraisal tools were specifically developed for other purposes, including clinical practice guideline development (theme 21), systematic review production (themes 22 and 23), and trial-reporting checklists (theme 24). Some

were based on frameworks for evidence-based practice (themes 25–27), which focus on trial design and interpretation evaluation to assess validity and applicability. Several articles mentioned unspecified structured tools, including homegrown tools, or use of an unspecified journal club approach (themes 28 and 29).

Tools were frequently referenced as potential methods for appraisal or as specifically used in studies and reviews of appraisal venues, such as journal clubs. Dimensions of appraisal were not frequently specified, but when they were, they typically focused on evidence validity and, to a lesser degree, applicability and reporting. No dimension focused on biological plausibility, ethics, or importance. In fact, these missing dimensions were highlighted as an important issue facing appraisal tools (theme 52). Similarly, it was discussed that there is a need for tailored tools for specific fields, problems, interventions, study designs, and outcomes (theme 53).

These structured, multidimensional tools were common for journal clubs. Many articles discussing settings other than journal clubs did not describe the use of tools. Some, however, proposed or described other approaches, including global ratings, such as simple numerical scores that are averages of subjective numerical user ratings (theme 30). Other ratings included automatically generated ones, such as journal impact factors (theme 31) and novel computational approaches aimed at identifying fraud (theme 32).

Category F: Appraisal venues

Appraisal of evidence occurs in a wide range of venues (category F). In-person group meetings as part of courses or journal clubs was the most common venue in our review (theme 33). Online forums were also common, specifically as a means for either performing a journal club as part of evidence-based practice or performing appraisal as part of the drive for more rigorous science (themes 34–38). Less commonly discussed were journal comments (theme 39) and simply reviewing the literature on one's own, outside of any specific setting (theme 40). Most unique was an institution's professional service that performed appraisal to answer evidentiary questions (theme 41).

Category G: Appraisal format

Regardless of the domains discussed or tools used, the appraisal process occurred in varied formats (category G). Journal clubs and course meetings were the most common, and were typically presentations or discussions attended synchronously by all attendees (theme 33). Many discussions were facilitated or moderated by either the presenter or another individual (theme 42). Less commonly, there were innovative formats, including debates (theme 43) and an approach that we call “generative-comparative,” whereby participants were prompted by the study question to propose an ideal study design to answer the question and to which they compared the actual study design (theme 44).

Online settings offered different approaches. Online journal clubs were typically asynchronous discussions where participants could comment at their convenience (theme 38). Several online platforms allow for indefinite posting, thereby enabling additional appraisals as information and contexts change (theme 45). Most online platforms utilized threaded comments (themes 35 and 37) or isolated comments (themes 34 and 36); however, online collaborative editing tools were also discussed as a format (theme 46).

Category H: Appraisal process issues

Several notable issues and concerns about the appraisal process were raised in the literature (category H). Some articles expressed concern

surrounding the possibility that appraisers, in an effort to not offend or affect future funding or publishing decisions, might be overly positive in their conclusions (theme 47). Conversely, some authors were concerned that appraisers, especially those in anonymous or online settings, might be more likely to state only negative criticisms (theme 47) or be uncivil (theme 48). Another concern was that appraisal, especially in open, anonymous forums, would result in comments that were of lower quality compared with those garnered by other venues, where appraisers might have more expertise or be more thoughtful (theme 49). Lastly, despite the potential benefits of novel online appraisal platforms, it was frequently mentioned that many of these tools are actually quite underused (theme 50).

Category I: Post-appraisal publication, organization, and use

After appraisal is performed, much can occur, or not, with the information generated (category I). Many articles discussing appraisal either specifically highlighted that the appraisal went unrecorded (theme 54) or did not specifically mention any recording, publication, or use outside of any learning that the appraiser experienced. Several studies did report that the appraisal content was recorded and locally published or stored for future local retrieval but not widely disseminated (theme 55).

When appraisals were disseminated, they were published either immediately online (theme 56) or within a journal (theme 57). Indexing or global aggregation of published appraisals was occasionally discussed (theme 58). Journal-based publications are automatically indexed within Medline; however, there was also discussion of indexing nonjournal published appraisals, though these discussions were theoretical. There was limited discussion of data standards (theme 59) and linkage to primary articles (theme 60).

There was some discussion about automatic responses to appraisals that are published. This included evaluating the quality of the appraisal (theme 61), messaging the appraisal to the authors of the original article (theme 62), having journal editors issue any needed corrections or retractions (theme 63), and having functionality for integrating appraisals related to a specific or set of dimensions (eg, research question, article author, appraiser, study design characteristic) (theme 64).

Category J: Post-appraisal concerns

Several issues were raised regarding appraisal data after publication (category J). Specifically, there were concerns that the volume of appraisals would worsen information overload (theme 65). Related to this concern is managing redundant appraisals (theme 66). Additionally, several articles highlighted the low quality of research on evidence appraisal (theme 67), particularly for journal clubs, where most of the research has been performed, as well as the low volume of research for both journal club and other formats of appraisal. Even when appraisals are documented, published, and accessible, they are underused (theme 68).

In sum, our scoping review demonstrates that evidence appraisal is occurring in a variety of different contexts (category F), by various stakeholders (category D), and to address a wide range of questions (themes 1–3 and 16). This appraisal occurs with varied tools (category E), formats (categories F and G), and approaches to dissemination (category I). Appraisal is insufficiently supported (themes 10 and 13) and researched (theme 67), lacks crucial infrastructure (themes 58–60), and is minimally applied (theme 68).

DISCUSSION

A conceptual framework

From our thematic analysis, we derived a conceptual framework for evidence appraisal (Figure 4). This was developed by first integrating emergent concepts related to evidence-appraisal processes and resources from the themes and theme categories, as well as domain knowledge in informatics. This was performed via iterative, consensus-driven discussion by the authors. Formal quantification of the significance and prevalence of these processes and resources is not within the scope of this review, but general levels of existence or gaps have been provided to the best of our knowledge as a result of performing this review and from our knowledge of relevant informatics resources.

Evidence appraisal is typically driven by a specific rationale, often related to a task, an institutional or professional requirement, or broader motivations, to ensure scientific rigor or pursue lifelong learning. Once convened, appraisal has several attributes. One attribute is the stage of the evidence's lifecycle. Appraisal can occur at the protocol stage, near the time of publication, or long after publication, when further contextual information is available, such as new standards of care or the identification of new adverse effects. Appraisal is also influenced by appraiser characteristics, tools used, venue, and format.

Once evidence has been appraised, the data from this appraisal instance can then be acquired as part of documentation or publication. The data can conform to data standards, and its capture can be aided by specific platforms or tools. The data can subsequently be mapped to an ontology, be indexed, or be curated and managed in other ways to improve data use. These processes help characterize and contextualize the data, essentially rendering it as information. This information can then be validated, integrated, and visualized, allowing for ultimate appraisal knowledge uptake as part of use cases. These use cases may include evidence synthesis, meta-research, scientific integrity assurance, methods development, determination of potential research questions, determination of research priorities, enabling of practice and policy decision-making, integration into education and training curricula, public science communication, and patient engagement. Our scoping review shows that this transformation from unrecorded content to actionable information and knowledge is underrealized.

Gaps and opportunities facing evidence appraisal

Research in this area is notably limited. No prior reviews of this topic have been performed. While there is significant research regarding journal clubs and PPPR, it is mainly descriptive and the trials were described as low quality. There are no Medical Subject Headings terms in PubMed for this field.

Our scoping review demonstrates that evidence appraisal faces significant challenges with data acquisition, management, organization, quality, coverage, availability, usability, and use. Underlying these issues is the lack of a significant organized field for evidence appraisal, and, as such, there is only minimal work to develop knowledge representation schemes, data standards, automatic knowledge acquisition or synthesis tools, and a central, aggregated, accessible, usable database for evidence appraisals. This has direct implications for data organization and usability, and may limit upstream data acquisition and coverage and downstream analysis and use. Enabling the transformation of appraisals from undocumented content to usable knowledge will require development of knowledge and technology tools, such as data standards, documentation tools, repositories, ontology, annotation tools, retrieval tools, validation systems, and integration and visualization systems.

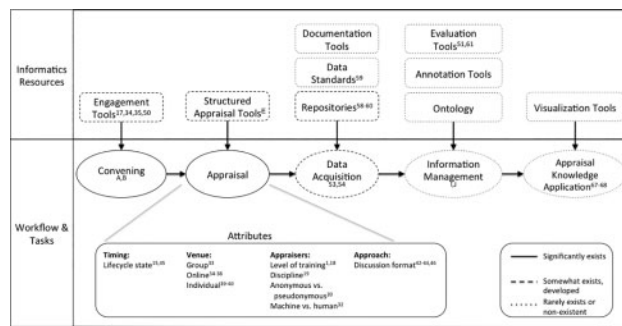


Figure 4. Conceptual framework. This figure illustrates the architecture of evidence appraisal as described or discussed in the biomedical literature and with actual or potential informatics resources to enable key steps. Superscript includes relevant themes or theme categories. The line style indicates whether these elements were described to exist or not within the scoping review or to the authors' knowledge.

While journal clubs appear to be ubiquitous, especially for medical trainees, they are learner-centric and lack data acquisition, limiting reuse of the appraisal content generated during discussions. Meanwhile, PPPR platforms, collaborative editing tools, and social media are increasingly discussed but appear to face underuse, fragmentation, and a lack of standardization. This discrepancy in appraisal production and use of data acquisition tools leads to downstream data issues. But this large volume of undocumented appraisals also reflects an opportunity, as appraisal production is already occurring at scale and only requires mechanisms and norms for data acquisition.

It is worth noting the rarity and absence of certain themes. There was little discussion of researchers or post-training practitioners performing journal clubs or of appraisal at academic conferences. There was little discussion of journal commentary, its evaluation, or the access barriers facing it. There was no discussion of developing approaches for lay stakeholders, such as patients and study participants, to appraise evidence, specifically related to study importance, ethical concerns, patient-centered outcome selection, or applicability factors.

Most important, there was minimal discussion regarding the establishment of a closed feedback loop whereby appraisals are not just generated by the research system, but are also utilized by it. Specifically, the research system could potentially utilize appraisal knowledge to improve public science communication, research prioritization, methods development, core outcome set determination, research synthesis, guideline development, and meta-research. Underlying the gap in appraisal uptake is the absence of institutions, working groups, and research funding mechanisms related to appraisal generation, processing, and dissemination.

An informatics research agenda

The key gaps facing evidence appraisal are related to increasing and improving the generation, acquisition, organization, integration, retrieval, and uptake of a large volume of appraisal data, which frequently involves complex natural language. Therefore, the evidence appraisal field would be best served by intensive research by the informatics community. Accordingly, we propose a specific research agenda informed by the knowledge of gaps identified in this review (Table 2). This agenda primarily focuses on related informatics methodology research, with novel application to the evidence appraisal domain. The agenda begins with enhancing appraisal data acquisition, organization, and integration by addressing gaps identified in the thematic analysis and conceptual framework, particularly the lack of en-

Table 2. An informatics research agenda

Informatics Research Opportunities	Subtopics
Develop data standards	Develop data and meta-data
Produce new datasets	Capture journal club and conference appraisal data
	Leverage electronic health record for data-driven appraisal
	Automate appraisal
Aggregate existing datasets	Published literature (eg, journal comments, reviews, meta-research, article introductions)
	Unindexed sources (eg, PPPR platforms, social media)
Develop appraisal ontology	Domain expert-driven
	Annotation-derived development
	Text-based concept detection
Expand tools and platforms	Researcher and practitioner engagement
	Patient and research participant engagement
	Retrieval tool development
	Recommendation system development
	Visualization tool development
Research appraisal knowledge	Methodology development
	Quality assessment
	Outcomes research
Enable other translational research	Study design, planning, and prioritization
	Trial methods development
	Novel application to nontraditional evidence sources
	Meta-research
	Systematic review development
	Practice/policy guideline development
Provide leadership and incentives	Working group formation
	Career, social, and financial incentivization for appraisal production
	Cultural norms for appraisal production and uptake
	Research funding
	Infrastructure development

agement tools, documentation tools, standards, aggregation, and repositories. This is to be accomplished via dataset aggregation, novel data production (including automated approaches), acquisition of data from current areas without documentation (eg, journal clubs), and development of data standards.

Next, we propose developing and applying an ontology for appraisal concepts and for data-quality and validity research. As noted in the conceptual framework, this area is, to our knowledge, nonexistent. Next, we propose developing tools and platforms that enable retrieval, recommendation, and visualization to enable appraisal knowledge application, which, again, is minimally existent outside of the task for which the appraisal was initially performed. Examples of potential appraisal concepts could include “trial arms lack equipoise,” “outcomes missing key patient-centered quality of life measure,” and “reported primary outcome is discrepant with the primary outcome described in the protocol.”

This will enable the next areas of research we propose, which also address the gap of adopting or applying appraisal knowledge: first, primary evidence appraisal knowledge research – descriptive, association-focused, and intervention research on appraisal information and knowledge resources; and second, use of appraisal information and knowledge to improve other translational research tasks via applied research. Primary appraisal research will include studying and developing appraisal methods, understanding relationships between clinical research and the appraisal of clinical research, and understanding relationships between appraisal data and scientific outcomes, such as representativeness, media sensationalization of science, science integrity issue detection, publication retraction, and clinical standard of care change. Enabled translational research uses of appraisal knowledge include integration during evidence synthesis, future study planning and prioritization, performance of meta-research, and the application of

appraisal to emerging and nontraditional sources of evidence, such as trial results in registries and repositories, journal articles lacking peer review, and analyses and conclusions in scientific blogging.

Lastly, sociotechnical work is an important component of our agenda. We propose that an evidence appraisal working group convene to better approach the aforementioned research, to promote funding and incentive changes, and to change cultural norms regarding the production and uptake of appraisal knowledge.

Limitations

This was a focused search that excluded synthesis tasks that might include appraisal components, did not perform snowball sampling, and had limited search terms due to high false positive rates. This approach allowed for feasible research with increased efficiency, but compromised the exhaustiveness of our results. Given our very broad research question and preference to include scientific discourse from other time periods, this was considered the most optimal approach. These limitations may affect references and themes identified. However, the goal of this review was not to conduct an exhaustive search or analysis, and was instead aimed at capturing common and emerging themes. Thematic analysis is inherently subjective, and other investigators may have arrived at different themes. Though it may be subject to these biases, this review was based on an open, iterative discourse to identify, modify, and ascribe themes.

CONCLUSIONS

While evidence appraisal has existed in a variety of forms for decades and has immense potential for enabling higher-value clinical research, it faces myriad obstacles. The field is fragmented, undefined,

and underresearched. Data from appraisal are rarely captured, organized, and transformed into usable knowledge. Appraisal knowledge is underutilized at other steps in the translational science pipeline. The evidence appraisal field lacks key research, infrastructure, and incentives. Despite these issues, discussion of appraisal is on the rise, and novel tools and data sources are emerging. We believe our proposed informatics research agenda provides a potential path forward to solidify and realize the potential of this emerging field.

ACKNOWLEDGMENTS

None.

FUNDING

This work was supported by R01 LM009886 (Bridging the Semantic Gap between Research Eligibility Criteria and Clinical Data; principal investigator, CW) and T15 LM007079 (Training in Biomedical Informatics at Columbia; principal investigator, Hripscak).

CONTRIBUTORS

AG proposed the methods, completed search and screening, performed thematic analysis, and drafted the manuscript. EV performed thematic analysis and contributed to the conceptual framework development and manuscript drafting. CW supervised the research, participated in study design, and edited the manuscript.

COMPETING INTERESTS

None.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

REFERENCES

- Chalmers I, Bracken MB, Djulbegovic B, *et al*. How to increase value and reduce waste when research priorities are set. *Lancet*. 2014;383(9912):156–65.
- Wooding S, Pollitt A, Castle-Clark S, *et al*. *Mental Health Retrosight: Understanding the Returns from Research (Lessons from Schizophrenia): Policy Report*. RAND Corporation; 2013. http://www.rand.org/content/dam/rand/pubs/research_reports/RR300/RR325/RAND_RR325.pdf. Accessed May 12, 2017.
- Wooding S, Hanney S, Buxton M, Grant J. *The Returns from Arthritis Research Volume 1: Approach, Analysis and Recommendations*. RAND Corporation; 2004. <http://www.rand.org/pubs/monographs/MG251.html>. Accessed May 12, 2017.
- Wooding S, Hanney S, Pollitt A, Buxton M, Grant J. *Project Retrosight: Understanding the returns from cardiovascular and stroke research: The Policy Report*. RAND Corporation; 2011. http://www.rand.org/pubs/working_papers/WR836.html. Accessed May 12, 2017.
- Chan AW, Song F, Vickers A, *et al*. Increasing value and reducing waste: addressing inaccessible research. *Lancet*. 2014;383(9913):257–66.
- Glasziou P, Altman DG, Bossuyt P, *et al*. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet*. 2014;383(9913):267–76.
- Ioannidis JP, Greenland S, Hlatky MA, *et al*. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet*. 2014;383(9912):166–75.
- Macleod MR, Michie S, Roberts I, *et al*. Biomedical research: increasing value, reducing waste. *Lancet*. 2014;383(9912):101–04.
- Deng G, Cassileth B. Complementary or alternative medicine in cancer care—myths and realities. *Nat Rev Clin Oncol*. 2013;10(11):656–64.
- Halpern SD, Karlawish JH, Berlin JA. The continuing unethical conduct of underpowered clinical trials. *JAMA*. 2002;288(3):358–62.
- Keen HI, Pile K, Hill CL. The prevalence of underpowered randomized clinical trials in rheumatology. *J Rheumatol*. 2005;32(11):2083–88.
- Savovic J, Jones HE, Altman DG, *et al*. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern Med*. 2012;157(6):429–38.
- Ocana A, Tannock IF. When are “positive” clinical trials in oncology truly positive? *J Natl Cancer Institute*. 2011;103(1):16–20.
- Kazi DS, Hlatky MA. Repeat revascularization is a faulty end point for clinical trials. *Circ Cardiovasc Qual Outcomes*. 2012;5(3):249–50.
- Ferreira-Gonzalez I, Busse JW, Heels-Ansdell D, *et al*. Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. *BMJ*. 2007;334(7597):786.
- Marcovitch H. Misconduct by researchers and authors. *Gac Sanit*. 2007;21(6):492–99.
- Gollogly L, Momen H. Ethical dilemmas in scientific publication: pitfalls and solutions for editors. *Rev Saude Publica*. 2006;40 (Spec no.):24–29.
- George SL, Buyse M. Data fraud in clinical trials. *Clin Investig (Lond)*. 2015;5(2):161–73.
- Tijdink JK, Verbeke R, Smulders YM. Publication pressure and scientific misconduct in medical scientists. *J Empir Res Hum Res Ethics*. 2014;9(5):64–71.
- Garcia-Berthou E, Alcaraz C. Incongruence between test statistics and P values in medical papers. *BMC Med Res Methodol*. 2004;4:13.
- Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, Williamson PR. Comparison of protocols and registry entries to published reports for randomised controlled trials. *Cochrane Database Syst Rev*. 2011;(1):MR000031.
- Gross C. Scientific Misconduct. *Annu Rev Psychol*. 2016;67:693–711.
- Altman DG. Poor-quality medical research: what can journals do? *JAMA*. 2002;287(21):2765–67.
- Altman DG. The scandal of poor medical research. *BMJ*. 1994;308(6924):283–84.
- Al-Shahi Salman R, Beller E, Kagan J, *et al*. Increasing value and reducing waste in biomedical research regulation and management. *Lancet*. 2014;383(9912):176–85.
- Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database: update and key issues. *New Engl J Med*. 2011;364(9):852–60.
- Dickersin K, Rennie D. The evolution of trial registries and their use to assess the clinical trial enterprise. *JAMA*. 2012;307(17):1861–64.
- Viergever RF, Ghersi D. The quality of registration of clinical trials. *PLoS One* 2011;6(2):e14701.
- Jule AM, Vaillant M, Lang TA, Guerin PJ, Olliaro PL. The schistosomiasis clinical trials landscape: a systematic review of antischistosomal treatment efficacy studies and a case for sharing individual participant-level data (IPD). *PLoS Negl Trop Dis*. 2016;10(6):e0004784.
- Tudur Smith C, Hopkins C, Sydes MR, *et al*. How should individual participant data (IPD) from publicly funded clinical trials be shared? *BMC Med*. 2015;13:298.
- Zarin DA, Tse T. Sharing Individual Participant Data (IPD) within the Context of the Trial Reporting System (TRS). *PLoS Med*. 2016;13(1):e1001946.
- Stahel PF, Moore EE. Peer review for biomedical publications: we can improve the system. *BMC Med*. 2014;12:179.
- van Roooyen S, Godlee F, Evans S, Black N, Smith R. Effect of open peer review on quality of reviews and on reviewers’ recommendations: a randomised trial. *BMJ*. 1999;318(7175):23–27.
- Ioannidis JP, Fanelli D, Dunne DD, Goodman SN. Meta-research: evaluation and improvement of research methods and practices. *PLoS Biol*. 2015;13(10):e1002264.

35. Kousta S, Ferguson C, Ganley E. Meta-research: broadening the scope of PLoS Biology. *PLoS Biol.* 2016;14(1):e1002334.
36. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ.* 2011;343:d5928.
37. Hrobjartsson A, Boutron I, Turner L, Altman DG, Moher D. Assessing risk of bias in randomised clinical trials included in Cochrane Reviews: the why is easy, the how is a challenge. *Cochrane Database Syst Rev.* 2013;(4):ED000058.
38. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence: study limitations (risk of bias). *J Clin Epidemiol.* 2011;64(4):407–15.
39. Ebrahim S, Sohani ZN, Montoya L, et al. Reanalyses of randomized clinical trial data. *JAMA.* 2014;312(10):1024–32.
40. Repeat after me. *Nat Med.* 2012;18(10):1443.
41. Landhuis E. Scientific literature: Information overload. *Nature.* 2016;535(7612):457–58.
42. Information overload. *Nature.* 2009;460(7255):551.
43. Ziman JM. The proliferation of scientific literature: a natural process. *Science.* 1980;208(4442):369–71.
44. Taichman DB, Backus J, Baethge C, et al. Sharing clinical trial data: a proposal from the International Committee of Medical Journal Editors. *Lancet.* 2016;387(10016):e9–11.
45. Smyth RL. Getting paediatric clinical trials published. *Lancet.* 2016;388(10058):2333–34.
46. Johnson AE, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data.* 2016;3:160035.
47. Hripscak G, Duke JD, Shah NH, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. *Stud Health Technol Inform.* 2015;216:574–78.
48. Kriegeskorte N. Open evaluation: a vision for entirely transparent post-publication peer review and rating for science. *Front Comput Neurosci.* 2012;6:79.
49. Kriegeskorte N, Walther A, Deca D. An emerging consensus for open evaluation: 18 visions for the future of scientific publishing. *Front Comput Neurosci.* 2012;6:94.
50. Yarkoni T. Designing next-generation platforms for evaluating scientific output: what scientists can learn from the social web. *Front Comput Neurosci.* 2012;6:72.
51. Sandefur CL. Young investigator perspectives. Blogging for electronic record keeping and collaborative research. *Am J Physiol Gastrointest Liver Physiol.* 2014;307(12):G1145–46.
52. Fausto S, Machado FA, Bento LF, Iamarino A, Nahas TR, Munger DS. Research blogging: indexing and registering the change in science 2.0. *PLoS One.* 2012;7(12):e51019.
53. Wolinsky H. More than a blog. Should science bloggers stick to popularizing science and fighting creationism, or does blogging have a wider role to play in the scientific discourse? *EMBO Rep.* 2011;12(11):1102–05.
54. Chatterjee P, Biswas T. Blogs and Twitter in medical publications: too unreliable to quote, or a change waiting to happen? *South African Med J.* 2011;101(10):712, 4.
55. Eyre-Walker A, Stoletzki N. The assessment of science: the relative merits of post-publication review, the impact factor, and the number of citations. *PLoS Biol.* 2013;11(10):e1001675.
56. Florian RV. Aggregating post-publication peer reviews and ratings. *Front Comput Neurosci.* 2012;6:31.
57. Galbraith DW. Redrawing the frontiers in the age of post-publication review. *Front Genet.* 2015;6:198.
58. Hunter J. Post-publication peer review: opening up scientific conversation. *Front Comput Neurosci.* 2012;6:63.
59. Knoepfler P. Reviewing post-publication peer review. *Trends Genet.* 2015;31(5):221–23.
60. Teixeira da Silva JA, Dobranszki J. Problems with traditional science publishing and finding a wider niche for post-publication peer review. *Accountability Res.* 2015;22(1):22–40.
61. Tierney E, O'Rourke C, Fenton JE. What is the role of 'the letter to the editor'? *Eur Arch Oto-rhino-laryngol.* 2015;272(9):2089–93.
62. Anderson KR. A new capability: postpublication peer review for pediatrics. *Pediatrics.* 1999;104(1 Pt 1):106.
63. Oxman AD, Sackett DL, Guyatt GH. Users' guides to the medical literature. I. How to get started. The Evidence-Based Medicine Working Group. *JAMA.* 1993;270(17):2093–95.
64. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA.* 1993;270(21):2598–601.
65. Committee to Review the Clinical and Translational Science Awards Program At the National Center for Advancing Translational Sciences. In: Leshner AI, Terry SF, Schultz AM, Liverman CT, eds. *The CTSA Program at NIH: Opportunities for Advancing Clinical and Translational Research.* Washington, DC; 2013.
66. Cortese DA, McGinnis JM. *Engineering a Learning Healthcare System: A Look at the Future: Workshop Summary.* Washington, DC; 2011.
67. Cortese DA, McGinnis JM. In: LA Olsen, DAisner, JMMcGinnis, eds. *The Learning Healthcare System: Workshop Summary.* Washington, DC; 2007.
68. Friedman CP, Wong AK, Blumenthal D. Achieving a nationwide learning health system. *Sci Transl Med.* 2010;2(57):57cm29.
69. Clancy C, Collins FS. Patient-Centered Outcomes Research Institute: the intersection of science and health care. *Sci Transl Med.* 2010;2(37):37cm18.
70. Gabriel SE, Normand SL. Getting the methods right: the foundation of patient-centered outcomes research. *New Engl J Med.* 2012;367(9):787–90.
71. Washington AE, Lipstein SH. The Patient-Centered Outcomes Research Institute: promoting better information, decisions, and health. *New Engl J Med.* 2011;365(15):e31.
72. McNutt M. Journals unite for reproducibility. *Science.* 2014;346(6210):679.
73. Collins FS, Tabak LA. Policy: NIH plans to enhance reproducibility. *Nature.* 2014;505(7485):612–13.
74. Landis SC, Amara SG, Asadullah K, et al. A call for transparent reporting to optimize the predictive value of preclinical research. *Nature.* 2012;490(7419):187–91.
75. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol.* 2005;8(1):19–32.
76. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 2009;6(7):e1000097.
77. Aiyer MK, Dorsch JL. The transformation of an EBM curriculum: a 10-year experience. *Med Teacher.* 2008;30(4):377–83.
78. Arif SA, Gim S, Nogid A, Shah B. Journal clubs during advanced pharmacy practice experiences to teach literature-evaluation skills. *Am J Pharmaceutical Educ.* 2012;76(5):88.
79. Balakas K, Sparks L. Teaching research and evidence-based practice using a service-learning approach. *J Nursing Educ.* 2010;49(12):691–95.
80. Biswas T. Role of journal clubs in undergraduate medical education. *Indian J Commun Med.* 2011;36(4):309–10.
81. Chakraborti C. Teaching evidence-based medicine using team-based learning in journal clubs. *Med Educ.* 2011;45(5):516–17.
82. Dawn S, Dominguez KD, Troutman WG, Bond R, Cone C. Instructional scaffolding to improve students' skills in evaluating clinical literature. *Am J Pharmaceutical Educ.* 2011;75(4):62.
83. Elnicki DM, Halperin AK, Shockcor WT, Aronoff SC. Multidisciplinary evidence-based medicine journal clubs: curriculum design and participants' reactions. *Am J Med Sci.* 1999;317(4):243–46.
84. Good DJ, McIntyre CM. Use of journal clubs within senior capstone courses: analysis of perceived gains in reviewing scientific literature. *J Nutr Educ Behav.* 2015;47(5):477–9.e1
85. Green ML. Evidence-based medicine training in graduate medical education: past, present and future. *J Eval Clin Pract.* 2000;6(2):121–38.
86. Horton R. Postpublication criticism and the shaping of clinical knowledge. *JAMA.* 2002;287(21):2843–47.

87. Oliver KB, Dalrymple P, Lehmann HP, McClellan DA, Robinson KA, Twose C. Bringing evidence to practice: a team approach to teaching skills required for an informationist role in evidence-based clinical and public health practice. *J Med Library Assoc.* 2008;96(1):50–57.
88. Smith-Strom H, Nortvedt MW. Evaluation of evidence-based methods used to teach nursing students to critically appraise evidence. *J Nursing Educ.* 2008;47(8):372–75.
89. Thompson CJ. Fostering skills for evidence-based practice: The student journal club. *Nurse Educ Pract.* 2006;6(2):69–77.
90. Willett LR, Kim S, Gochfeld M. Enlivening journal clubs using a modified 'jigsaw' method. *Med Educ.* 2013;47(11):1127–28.
91. Alguire PC. A review of journal clubs in postgraduate medical education. *J General Int Med.* 1998;13(5):347–53.x
92. Austin TM, Richter RR, Frese T. Using a partnership between academic faculty and a physical therapist liaison to develop a framework for an evidence-based journal club: a discussion. *Physiotherapy Res Int.* 2009;14(4):213–23.
93. Berger J, Hardin HK, Topp R. Implementing a virtual journal club in a clinical nursing setting. *J Nurses Staff Dev.* 2011;27(3):116–20.
94. Campbell-Fleming J, Catania K, Courtney L. Promoting evidence-based practice through a traveling journal club. *Clin Nurse Specialist.* 2009;23(1):16–20.
95. Crank-Patton A, Fisher JB, Toedter LJ. The role of the journal club in surgical residency programs: a survey of APDS program directors. *Current Surgery.* 2001;58(1):101–04.
96. Deenadayalan Y, Grimmer-Somers K, Prior M, Kumar S. How to run an effective journal club: a systematic review. *J Eval Clin Pract.* 2008;14(5):898–911.
97. Dirschl DR, Tornetta P 3rd, Bhandari M. Designing, conducting, and evaluating journal clubs in orthopaedic surgery. *Clin Orthopaedics Related Res.* 2003;413:146–57.
98. Doust J, Del Mar CB, Montgomery BD, et al. EBM journal clubs in general practice. *Australian Family Phys.* 2008;37(1–2):54–56.
99. Dovi G. Empowering change with traditional or virtual journal clubs. *Nurs Manag.* 2015;46(1):46–50.
100. Duffy JR, Thompson D, Hobbs T, Niemeyer-Hackett NL, Elpers S. Evidence-based nursing leadership: Evaluation of a Joint Academic-Service Journal Club. *J Nursing Admin.* 2011;41(10):422–27.
101. Ebbert JO, Montori VM, Schultz HJ. The journal club in postgraduate medical education: a systematic review. *Med Teacher.* 2001;23(5):455–61.
102. Figueroa RA, Valdivieso S, Turpaud M, Cortes P, Barros J, Castano C. Journal club experience in a postgraduate psychiatry program in Chile. *Acad Psychiatr.* 2009;33 (5):407–09.
103. Forsen JW Jr, Hartman JM, Neely JG. Tutorials in clinical research, part VIII: creating a journal club. *The Laryngoscope.* 2003;113(3):475–83.
104. Grant MJ. Journal clubs for continued professional development. *Health Inform Libraries J.* 2003;20 (Suppl 1):72–73.
105. Heiligman RM, Wollitzer AO. A survey of journal clubs in U.S. family practice residencies. *J Med Educ.* 1987;62(11):928–31.
106. Hinkson CR, Kaur N, Sipes MW, Pierson DJ. Impact of offering continuing respiratory care education credit hours on staff participation in a respiratory care journal club. *Respir Care.* 2011;56(3):303–05.
107. Hunt C, Topham L. Setting up a multidisciplinary journal club in learning disability. *Brit J Nurs (Mark Allen Publishing).* 2002;11(10):688–93.
108. Jones SR, Harrison MM, Crawford IW, et al. Journal clubs in clinical medicine. *Emerg Med J.* 2002;19(2):184–85.
109. Kelly AM, Cronin P. Setting up, maintaining and evaluating an evidence based radiology journal club: the University of Michigan experience. *Acad Radiol.* 2010;17(9):1073–78.
110. Khan KS, Dwarakanath LS, Pakkal M, Brace V, Awonuga A. Postgraduate journal club as a means of promoting evidence-based obstetrics and gynaecology. *J Obstetrics Gynaecol.* 1999;19(3):231–34.
111. Lachance C. Nursing journal clubs: a literature review on the effective teaching strategy for continuing education and evidence-based practice. *J Continuing Educ Nurs.* 2014;45(12):559–65.
112. Leung EY, Malick SM, Khan KS. On-the-job evidence-based medicine training for clinician-scientists of the next generation. *Clin Biochem Rev.* 2013;34(2):93–103.
113. Linzer M. The journal club and medical education: over one hundred years of unrecorded history. *Postgraduate Med J.* 1987;63(740):475–78.
114. Lizarondo L, Kumar S, Grimmer-Somers K. Online journal clubs: an innovative approach to achieving evidence-based practice. *J Allied Health.* 2010;39(1):e17–22.
115. Lizarondo LM, Grimmer-Somers K, Kumar S. Exploring the perspectives of allied health practitioners toward the use of journal clubs as a medium for promoting evidence-based practice: a qualitative study. *BMC Med Educ.* 2011;11:66.
116. Luby M, Riley JK, Towne G. Nursing research journal clubs: bridging the gap between practice and research. *Medsurg Nurs.* 2006;15(2):100–02.
117. Matthews DC. Journal clubs most effective if tailored to learner needs. *Evid Based Dentistry.* 2011;12(3):92–93.
118. Miller SA, Forrest JL. Translating evidence-based decision making into practice: appraising and applying the evidence. *J Evid Based Dent Pract.* 2009;9(4):164–82.
119. Mobbs RJ. The importance of the journal club for neurosurgical trainees. *J Clin Neurosci.* 2004;11(1):57–58.
120. Moro JK, Bhandari M. Planning and executing orthopedic journal clubs. *Indian J Orthopaedics.* 2007;41(1):47–54.
121. Pato MT, Cobb RT, Lusskin SI, Schardt C. Journal club for faculty or residents: a model for lifelong learning and maintenance of certification. *Int Rev Psychiatry (Abingdon, England).* 2013;25(3):276–83.
122. Pearce-Smith N. A journal club is an effective tool for assisting librarians in the practice of evidence-based librarianship: a case study. *Health Inform Libraries J* 2006;23(1):32–40.
123. Price DW, Felix KG. Journal clubs and case conferences: from academic tradition to communities of practice. *J Continuing Educ Health Professions.* 2008;28(3):123–30.
124. Rheeder P, van Zyl D, Webb E, Worku Z. Journal clubs: knowledge sufficient for critical appraisal of the literature? *South African Med J.* 2007;97(3):177–78.
125. Seelig CB. Affecting residents' literature reading attitudes, behaviors, and knowledge through a journal club intervention. *J General Int Med.* 1991;6(4):330–34.
126. Seymour B, Kinn S, Sutherland N. Valuing both critical and creative thinking in clinical practice: narrowing the research-practice gap? *J Advan Nurs.* 2003;42(3):288–96.
127. Shannon S. Critically appraised topics (CATs). *Can Assoc Radiol J.* 2001;52(5):286–87.
128. Sheehan J. A journal club as a teaching and learning strategy in nurse teacher education. *J Advan Nurs.* 1994;19(3):572–78.
129. Stewart C, Snyder K, Sullivan SC. Journal clubs on the night shift: a staff nurse initiative. *Medsurg Nurs.* 2010;19(5):305–06.
130. Topf JM, Sparks MA, Iannuzzella F, et al. Twitter-based journal clubs: additional facts and clarifications. *J Med Internet Res.* 2015;17(9):e216.
131. Westlake C, Albert NM, Rice KL, et al. Nursing journal clubs and the clinical nurse specialist. *Clin Nurse Special.* 2015;29(1):E1–e10.
132. Clark E, Burkett K, Stanko-Lopp D. Let Evidence Guide Every New Decision (LEGEND): an evidence evaluation system for point-of-care clinicians and guideline development teams. *J Eval Clin Pract.* 2009;15(6):1054–60.
133. Schunemann HJ, Bone L. Evidence-based orthopaedics: a primer. *Clin Orthopaedics Related Res.* 2003;413:117–32.
134. Bachmann T. Fair and open evaluation may call for temporarily hidden authorship, caution when counting the votes, and transparency of the full pre-publication procedure. *Front Comput Neurosci.* 2011;5:61.
135. Bellomo R, Bagshaw SM. Evidence-based medicine: classifying the evidence from clinical trials – the need to consider other dimensions. *Critical Care (London, England).* 2006;10(5):232.
136. Dreier M, Borutta B, Stahmeyer J, Krauth C, Walter U. Comparison of tools for assessing the methodological quality of primary and secondary

- studies in health technology assessment reports in Germany. *GMS Health Technol Assess*. 2010;6:Doc07.
137. Ghosh SS, Klein A, Avants B, Millman KJ. Learning from open source software projects to improve scientific review. *Front Comput Neurosci*. 2012;6:18.
 138. Poschl U. Multi-stage open peer review: scientific evaluation integrating the strengths of traditional peer review with the virtues of transparency and self-regulation. *Front Comput Neurosci*. 2012;6:33.
 139. Woolf SH. Taking critical appraisal to extremes. The need for balance in the evaluation of evidence. *J Fam Pract*. 2000;49(12):1081–85.
 140. Wright J. Journal clubs: science as conversation. *New Engl J Med*. 2004;351(1):10–12.
 141. Ahmadi N, McKenzie ME, Maclean A, Brown CJ, Mastracci T, McLeod RS. Teaching evidence based medicine to surgery residents: is journal club the best format? A systematic review of the literature. *J Surg Educ*. 2012;69(1):91–100.
 142. Akhund S, Kadir MM. Do community medicine residency trainees learn through journal club? An experience from a developing country. *BMC Med Educ*. 2006;6:43.
 143. Bednarczyk J, Pauls M, Fridfinnson J, Weldon E. Characteristics of evidence-based medicine training in Royal College of Physicians and Surgeons of Canada emergency medicine residencies – a national survey of program directors. *BMC Med Educ*. 2014;14:57.
 144. Grant WD. An evidence-based journal club for dental residents in a GPR program. *J Dental Educ*. 2005;69(6):681–86.
 145. Greene WB. The role of journal clubs in orthopaedic surgery residency programs. *Clin Orthopaedics Related Res*. 2000;373:304–10.
 146. Ibrahim A, Mshelbwala PM, Mai A, Asuku ME, Mbibu HN. Perceived role of the journal clubs in teaching critical appraisal skills: a survey of surgical trainees in Nigeria. *Nigerian J Surgery*. 2014;20(2):64–68.
 147. Jouriles NJ, Cordell WH, Martin DR, Wolfe R, Emerman CL, Avery A. Emergency medicine journal clubs. *Acad Emerg Med*. 1996;3(9):872–78.
 148. Laibhen-Parkes N. Increasing the practice of questioning among pediatric nurses: “The Growing Culture of Clinical Inquiry” project. *J Pediatric Nurs*. 2014;29(2):132–42.
 149. Landi M, Springer S, Estus E, Ward K. The impact of a student-run journal club on pharmacy students’ self-assessment of critical appraisal skills. *Consultant Pharmacist*. 2015;30(6):356–60.
 150. Lloyd G. Journal clubs. *J Accident Emerg Med*. 1999;16(3):238–39.
 151. Mattingly D. Proceedings of the conference on the postgraduate medical centre. *Journal clubs*. *Postgraduate Med J*. 1966;42(484):120–22.
 152. Mazuryk M, Daeninck P, Neumann CM, Bruera E. Daily journal club: an education tool in palliative care. *Palliative Med*. 2002;16(1):57–61.
 153. Millichap JJ, Goldstein JL. Neurology Journal Club: a new subsection. *Neurology*. 2011;77(9):915–17.
 154. Moharari RS, Rahimi E, Najafi A, Khashayar P, Khajavi MR, Meysamie AP. Teaching critical appraisal and statistics in anesthesia journal club. *QJM*. 2009;102(2):139–41.
 155. Phillips RS, Glasziou P. What makes evidence-based journal clubs succeed? *ACP J Club*. 2004;140(3):A11–12.
 156. Phitayakorn R, Gelula MH, Malangoni MA. Surgical journal clubs: a bridge connecting experiential learning theory to clinical practice. *J Am College Surgeons*. 2007;204(1):158–63.
 157. Pitner ND, Fox CA, Riess ML. Implementing a successful journal club in an anesthesiology residency program. *F1000Research*. 2013;2:15.
 158. Serghi A, Goebert DA, Andrade NN, Hishinuma ES, Lunsford RM, Matsuda NM. One model of residency journal clubs with multifaceted support. *Teaching Learning Med*. 2015;27(3):329–40.
 159. Shokouhi G, Ghojzadeh M, Sattarnezhad N. Organizing Evidence Based Medicine (EBM) Journal Clubs in Department of Neurosurgery, Tabriz University of Medical Sciences. *Int J Health Sci*. 2012;6(1):59–62.
 160. Stern P. Using journal clubs to promote skills for evidence-based practice. *Occup Therapy Health Care*. 2008;22(4):36–53.
 161. Tam KW, Tsai LW, Wu CC, Wei PL, Wei CF, Chen SC. Using vote cards to encourage active participation and to improve critical appraisal skills in evidence-based medicine journal clubs. *J Eval Clin Pract*. 2011;17(4):827–31.
 162. Temple CL, Ross DC. Acquisition of evidence-based surgery skills in plastic surgery residency training. *J Surg Educ*. 2011;68(3):167–71.
 163. Thangasamy IA, Leveridge M, Davies BJ, Finelli A, Stork B, Woo HH. International Urology Journal Club via Twitter: 12-month experience. *Eur Urol*. 2014;66(1):112–17.
 164. Wilson M, Ice S, Nakashima CY, et al. Striving for evidence-based practice innovations through a hybrid model journal club: A pilot study. *Nurse Educ Today*. 2015;35(5):657–62.
 165. Young T, Rohwer A, Volmink J, Clarke M. What are the effects of teaching evidence-based health care (EBHC)? Overview of systematic reviews. *PLoS One*. 2014;9(1):e86706.
 166. Misra UK, Kalita J, Nair PP. Traditional journal club: a continuing problem. *J Assoc Phys India*. 2007;55:343–46.
 167. Sidorov J. How are internal medicine residency journal clubs organized, and what makes them successful? *Arch Int Med*. 1995;155(11):1193–97.
 168. Walther A, van den Bosch JJ. FOSE: a framework for open science evaluation. *Front Comput Neurosci*. 2012;6:32.
 169. Del Mar CB, Glasziou PP. Ways of using evidence-based medicine in general practice. *Med J Australia*. 2001;174(7):347–50.
 170. Tracy SL. From bench-top to chair-side: how scientific evidence is incorporated into clinical practice. *Dental Mat*. 2014;30(1):1–15.
 171. Bowden SC, Harrison EJ, Loring DW. Evaluating research for clinical significance: using critically appraised topics to enhance evidence-based neuropsychology. *Clin Neuropsychol*. 2014;28(4):653–68.
 172. Fetters L, Figueiredo EM, Keane-Miller D, McSweeney DJ, Tsao CC. Critically appraised topics. *Pediatric Phys Therapy*. 2004;16(1):19–21.
 173. Kelly AM, Cronin P. How to perform a critically appraised topic: part 2, appraise, evaluate, generate, and recommend. *Am J Roentgenol*. 2011;197(5):1048–55.
 174. Faulkes Z. The vacuum shouts back: postpublication peer review on social media. *Neuron*. 2014;82(2):258–60.
 175. Rychetnik L, Frommer M, Hawe P, Shiell A. Criteria for evaluating evidence on public health interventions. *J Epidemiol Commun Health*. 2002;56(2):119–27.
 176. Burchett H, Umoquit M, Dobrow M. How do we know when research from one setting can be useful in another? A review of external validity, applicability and transferability frameworks. *J Health Services Res Policy*. 2011;16(4):238–44.
 177. Leung EY, Tirlapur SA, Siassakos D, Khan KS. #BlueJC: BJOG and Katherine Twining Network collaborate to facilitate post-publication peer review and enhance research literacy via a Twitter journal club. *BJOG*. 2013;120(6):657–60.
 178. Crowe M, Sheppard L, Campbell A. Comparison of the effects of using the Crowe Critical Appraisal Tool versus informal appraisal in assessing health research: a randomised trial. *Int J Evid Based Healthcare*. 2011;9(4):444–49.
 179. Gonzalez LS 3rd. Referees make journal clubs fun. *BMJ (Clinical research ed)*. 2003;326(7380):106.
 180. Manning ML, Davis J. Journal Club: a venue to advance evidence-based infection prevention practice. *Am J Infect Control*. 2012;40(7):667–69.
 181. Coomarasamy A, Taylor R, Khan KS. A systematic review of postgraduate teaching in evidence-based medicine and critical appraisal. *Med Teacher*. 2003;25(1):77–81.
 182. Collier R. When postpublication peer review stings. *Can Med Assoc J*. 2014;186(12):904.
 183. Perera M, Roberts M, Lawrentschuk N, Bolton D. Response to “Twitter-Based Journal Clubs: Some Additional Facts and Clarifications.” *J Med Internet Res*. 2015;17(9):e217.
 184. Roberts MJ, Perera M, Lawrentschuk N, Romanic D, Papa N, Bolton D. Globalization of continuing professional development by journal clubs via microblogging: a systematic review. *J Med Internet Res*. 2015;17(4):e103.
 185. Topf JM, Hiremath S. Social media, medicine and the modern journal club. *Int Rev Psychiatry (Abingdon, England)*. 2015;27(2):147–54.
 186. Slavov N. Making the most of peer review. *eLife*. 2015;4:e12708.