

**PERSPECTIVE**

# Time to challenge the spurious hierarchy of systematic over narrative reviews?

Trisha Greenhalgh<sup>1</sup>  | Sally Thorne<sup>2</sup>  | Kirsti Malterud<sup>3</sup> <sup>1</sup>Department of Primary Care Health Sciences, University of Oxford, Oxford, UK<sup>2</sup>School of Nursing, University of British Columbia, Vancouver, Canada<sup>3</sup>Research Unit for General Practice, Uni Research Health, Bergen, Norway**Correspondence**

Trisha Greenhalgh, Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK.

Email: trish.greenhalgh@phc.ox.ac.uk

**Funding information**

TG's contribution was part-funded from the National Institute for Health Research Biomedical Research Centre, Oxford, Grant BRC-1215-20008.

## 1 | BACKGROUND

Cynthia Mulrow's important paper calling for literature reviews to be undertaken more systematically (and hence be more informative and reliable) is now 30 years old.<sup>1</sup> A recent paper in *BMC Medical Research Methodology* compared the proportion of reviews that were systematic (as opposed to narrative) in five leading biomedical journals.<sup>2</sup> The authors found significant diversity: from *New England Journal of Medicine* (0%) and *Lancet* (11%) to *Annals of Internal Medicine* (72%). *Systematic reviews* were assumed by the authors to be superior because they are (i) more likely to have a focused research question, (ii) more methodologically explicit and (iii) less likely to be biased than *narrative reviews*.

This stance reflects the *raison d'être* of the Cochrane Collaboration, whose use of explicit and auditable quality criteria for undertaking systematic reviews has inspired a weighty methodological handbook,<sup>3</sup> numerous tools and checklists<sup>4,5</sup> and structured reporting criteria.<sup>6</sup> There is strong emphasis on methodological reproducibility, with the implication that a different review team, using the same search criteria, quality checklists and synthesis tools, should obtain the same result.<sup>3</sup>

Yet leading medical journals regularly publish clinical topic reviews that may lack a focused research question,

methods section or statement on how studies were selected and analysed (see for example<sup>7-9</sup>). These narrative reviews typically draw on expert opinion by deliberately recruiting leading names in the field (eg "The aim of this Commission is to provide the strongest evidence base *through involvement of experts* from a wide cross-section of disciplines..."—page 1953, emphasis added<sup>8</sup>). Reviews crafted through the experience and judgement of experts are often viewed as untrustworthy ("eminence-based" is a pejorative term). Yet the classical definition of the EBM as "the conscientious, explicit, and *judicious* use of current best evidence..." (page 71, emphasis added)<sup>10</sup> suggests a key role for judgement in the selection and interpretation of evidence.

In short, there appears to be a growing divergence between the assumed "hierarchy" of evidence in secondary research, which defines systematic reviews as superior,<sup>11</sup> and what some leading academic journals view as a state-of-the-art (that is, expert-led narrative) review. We believe this is partly because the systematic review format has been erroneously defined as a universal gold standard and partly because the term "narrative review" is frequently misunderstood, misapplied and unfairly dismissed.

Systematic reviews in the Cochrane sense use a highly technical approach to identification, appraisal and synthesis of evidence and typically (although not invariably)

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2018 The Authors. *European Journal of Clinical Investigation* published by John Wiley & Sons Ltd on behalf of Stichting European Society for Clinical Investigation Journal Foundation.

privilege randomised controlled trials or previous systematic reviews over other forms of evidence.<sup>11</sup> This may be entirely appropriate—especially when the primary purpose is to answer a very specific question about how to treat a particular disease in a particular target group.

But the doctor in the clinic, the nurse on the ward or the social worker in the community will encounter patients with a wide diversity of health states, cultural backgrounds, illnesses, sufferings and resources.<sup>12</sup> And those who gather around the policymaking table will find multiple calls on their attention—including burden of need, local availability of different treatments, personal testimony, strength of public opinion and budgetary realities. To produce a meaningful synthesis of research evidence relevant to such complex situations, the reviewer must (i) incorporate a broad range of knowledge sources and strategies for knowing and (ii) undertake multi-level interpretation using creativity and judgement.<sup>12,13</sup>

We align with previous authors, who, drawing on Wittgenstein, distinguish between puzzles or problems that require *data* (for which a conventional systematic review, with meta-analysis where appropriate, may be the preferred methodology) and those that require *clarification and insight* (for which a more interpretive and discursive synthesis of existing literature is needed).<sup>14,15</sup>

Below, we explore both strengths, limitations and conceptual confusions of systematic and narrative reviews. We consider three questions: what makes a review systematic; what is a narrative review and whether these different kinds of review should be viewed as competing or complementary.

## 2 | WHAT MAKES A REVIEW SYSTEMATIC?

The defining characteristic of a systematic review in the Cochrane sense is the use of a predetermined structured method to search, screen, select, appraise and summarise study findings to answer a narrowly focused research question.<sup>3,16</sup> Using an exhaustive search methodology, the reviewer extracts all possibly relevant primary studies, and then limits the dataset using explicit inclusion and exclusion criteria. The review focus is highly circumscribed and quality criteria are tightly enforced. Typically, a body of hundreds or thousands of potential studies identified in the initial search is whittled down to a mere handful before the reviewer even begins to consider what they collectively mean.

The term “systematic” is thus by no means synonymous with “high-quality”. Rather, it can be viewed as a set of methodologies characterised by tight focus, exhaustive search, high rejection-to-inclusion ratio and an emphasis on technical rather than interpretive synthesis methods.

The conflation of the quality of a review with the assiduousness of such tasks as searching, applying inclusion and

### Key point

- Systematic reviews are generally placed above narrative reviews in an assumed hierarchy of secondary research evidence
- We argue that systematic reviews and narrative reviews serve different purposes and should be viewed as complementary
- Conventional systematic reviews address narrowly focused questions; their key contribution is summarising data
- Narrative reviews provide interpretation and critique; their key contribution is deepening understanding

exclusion criteria, creating tables of extracted data and mathematically summing effect sizes (rather than, for example, with the level of critical analysis of the papers’ unstated assumptions and discussion sections) has, we believe, led to a proliferation of systematic reviews that represent aggregations of findings within the narrow body of work that has met the authors’ eligibility criteria.<sup>17-19</sup> Such studies may sometimes add value, especially when additional meta-analysis confirms whether a clinically significant effect is or is not also statistically significant.<sup>20</sup> But sometimes, the term “systematic review” allows a data aggregation to claim a more privileged position within the knowledge hierarchy than it actually deserves.<sup>11</sup>

We acknowledge that the science of systematic review within the Cochrane and Campbell Collaborations is evolving to embrace a wider range of primary studies and methodologies, with recommended procedures for sampling, assessment and synthesis of evidence compliant with the question asked and the context explored. The adjective “systematic” is thus coming to acquire a broader meaning in terms of the transparency and appropriateness of methods, rather than signifying strict adherence to a particular pre-defined tool or checklist or a privileging of randomised trials (see for example methodological work by Lewin et al,<sup>21</sup> Petticrew et al<sup>22</sup> and Pluye et al<sup>23-25</sup>). All these approaches, however, remain focused on answering a relatively narrow question that is pre-defined at the outset and with a primary focus on the extraction, tabulation and summation of empirical data.

## 3 | WHAT IS A NARRATIVE REVIEW?

A *narrative review* is a scholarly summary along with interpretation and critique.<sup>26</sup> It can be conducted using a number of distinctive methodologies. While principles and

procedures may diverge from the classic methodology of systematic review, they are not unsystematic (in the sense of being *ad hoc* or careless), and may certainly be conducted and presented in a systematic way, depending on purpose, method and context.

Different kinds of reviews offer different kinds of truth: the conventional systematic review with meta-analysis deals in *probabilistic* (typically, Bayesian) truth; it is concerned mainly with producing generalisable “facts” to aid prediction. The narrative review, in contrast, deals in *plausible* truth. Its goal is an authoritative argument, based on informed wisdom that is convincing to an audience of fellow experts. To that end, the author of a narrative review must authentically represent in the written product both the underpinning evidence (including but not limited to primary research) and *how* this evidence has been drawn upon and drawn together to inform the review’s conclusions.

A *hermeneutic review* takes as its reference point the notion of *verstehen*, or the process of creating an interpretive understanding.<sup>14</sup> It capitalises on the continual deepening of insight that can be obtained by critical reflection on particular elements of a dataset—in this case, individual primary studies—in the context of a wider body of work. It may or may not define its reference body of studies using systematic search methods and inclusion/exclusion criteria, but its primary focus is on the essential tasks of induction and interpretation in relation to the defined sample for the purpose of advancing theoretical understanding.<sup>17</sup> A *realist review* considers the “generative causality,” in which particular mechanisms (for example, peer influence) produce particular outcomes (for example, smoking cessation) in some circumstances (for example, when societal disapproval of smoking is high) but not others (for example, in cultures where smoking is still widely viewed as a mark of sophistication).<sup>27</sup> A *meta-narrative review* maps the storyline of a research tradition over time.<sup>28</sup> Shifting the focus away from comparing findings of studies published at different times, it orients critical reflection to discern how ideas have waxed and waned within different scholarly communities at different points in the development of thinking (see an early example of how the term “diffusion of innovations” was differently defined and explored in different academic disciplines<sup>29</sup>).

Each of these forms of narrative review (along with other specialist approaches to combining primary studies in qualitative research<sup>30,31</sup>) reflects an explicit lens that is expected to shape the understandings that will arise from the review process, through analysis and synthesis processes that may be highly systematic. Narrative reviews also include a number of more generic styles such as integrative<sup>32,33</sup> and critical,<sup>34</sup> the former being the approach generally taken by narrative reviews in clinical journals. All these approaches play an important role in expanding our

understanding not only of the topic in question but also of the reasons why it has been studied in a particular way, the interpretations that have been variously made with respect to what we know about it, and the nature of the knowledge base that informs or might inform clinical practice.

Because hermeneutic, realist and meta-narrative reviews have explicit methodologies and accepted standards and criteria for judging their quality,<sup>14,27,28</sup> a minority of scholars include such approaches within the (broadly defined) category of systematic reviews. However, we have had experience of journal editors rejecting reviews based on these techniques on the grounds that they were “not systematic”. Also of note is the emergence of “how-to” guides for narrative reviews, which (misleadingly in our view) exhort the reviewer to focus carefully on such tasks as starting with an explicit search strategy and defining strict inclusion and exclusion criteria for primary studies.<sup>35,36</sup> In other words, the boundaries between systematic and narrative reviews are both fuzzy and contested.

#### 4 | SYSTEMATIC OR NARRATIVE OR SYSTEMATIC AND NARRATIVE?

The conflation of “systematic” with superior quality (and “narrative” with inferior quality) has played a major role in the muddying of methodological waters in secondary research. This implicit evidence hierarchy (or pyramid) elevates the mechanistic processes of exhaustive search, wide exclusion and mathematical averaging over the thoughtful, in-depth, critically reflective processes of engagement with ideas. The emphasis on *thinking* and *interpretation* in narrative review has prompted some authors to use the term “evidence-informed” rather than “evidence-based”<sup>15,37</sup>: the narrative review is both less and more than a methods-driven exercise in extracting and summing data.

Training in systematic reviews has produced a generation of scholars who are skilled in the technical tasks of searching, sorting, checking against inclusion criteria, tabulating extracted data and generating “grand means” and confidence intervals.<sup>3</sup> These skills are important, but as the recent article by Faggion et al illustrates, critics may incorrectly assume that they override and make redundant the generation of understanding. To the extent that the term “systematic review” privileges only that which is common in the findings amongst a rigidly defined subset of the available body of work, we risk losing sight of the marvellous diversities and variations that ought to intrigue us. In excluding those aspects of scholarship, systematic reviews hold the potential to significantly skew our knowledge landscape. While there are occasions when systematic review is the ideal approach to answering specific types of

question, the absence of thoughtful, interpretive critical reflection can render such products hollow, misleading and potentially harmful.

The argument that systematic reviews are less biased than narrative reviews begs the question of what we mean by bias. *Bias* is an epidemiological construct, which refers to something that distorts the objective comparisons between groups.<sup>20</sup> It presupposes the dispassionate, instrumental and universal “view from nowhere” that has long defined the scientific method.<sup>38</sup> When we speak of interpretation, we refer to an analysis that is necessarily perspectival, with the interpreter transparently positioned in order that the reader can understand *why* this particular perspective, selection process and interpretive methodology was selected in relation to the question at hand.<sup>14,17,29,37,39</sup> Systematic and transparent reflection upon and sharing of such aspects of the research process adds to the scientific quality of interpretive research.

The question of whether “systematic” review techniques can eliminate bias in secondary research is in any case questionable. The privileging of freedom from bias over *relevance* of question and findings wrongly assumes that how the topic is framed, and which questions should be explored is somehow self-evident. A recent review of systematic reviews generated by a national knowledge centre to inform policymaking in Norway showed that in most cases, the evidence base addressed only a fraction of relevant policy questions.<sup>40</sup> More generally, there is growing evidence that the science of systematic reviews is becoming increasingly distorted by commercial and other conflicts of interest, leading to reviews, which—often *despite* ticking the boxes on various quality checklists—are unnecessary, misleading or partisan.<sup>19,41</sup> The holy grail of a comprehensive database of unambiguous and unbiased evidence summaries (in pursuit of which the Cochrane Collaboration was founded<sup>42</sup>) continues to recede into the future.

A legitimate criticism of narrative reviews is that they may “cherry pick” evidence to bolster a particular perspective. But this must be weighed against the counter-argument that the narrative reviewer selects evidence judiciously and purposively with an eye to what is relevant for key policy questions—including the question of which future research programmes should be funded. Whilst we accept that narrative reviews can be performed well or badly, we believe the undervaluing of such reviews is a major contributor to research waste. In the absence of an interpretive overview of a topic that clearly highlights the state of knowledge, ignorance and uncertainty (explaining how we know what we know, and where the intriguing unanswered questions lie), research funding will continue to be ploughed into questions that are of limited importance, and which have often already been answered.<sup>40</sup>

This principle was illustrated in a recent hermeneutic review of telehealth in heart failure by one of us.<sup>43</sup> It identified 7 systematic reviews of systematic reviews, 32 systematic reviews (including 17 meta-analyses) covering hundreds of primary studies, as well as six mega-trials—almost all of which had concluded that more research (addressing the same narrow question with yet more randomised trials intended to establish an effect size for telehealth) was needed. The hermeneutic approach revealed numerous questions that had remained under-explored as researchers had pursued this narrow question—including the complex and changing nature of the co-morbidities and social determinants associated with heart failure, the varied experiences and priorities of patients with heart failure, the questionable nature of up-titration as a guiding principle in heart failure management, and the numerous organisational, regulatory and policy-level complexities associated with introducing telehealth programmes. The review concluded that: “The limited adoption of telehealth for heart failure has complex clinical, professional and institutional causes, which are unlikely to be elucidated by adding more randomised trials of technology-on versus technology-off to an already-crowded literature. An alternative approach is proposed, based on naturalistic study designs, application of social and organisational theory, and co-design of new service models based on socio-technical principles” (page 156).

## 5 | CONCLUSION

As many authors and journal editors are well aware, the narrative review is not a poor cousin of the systematic review but a different and potentially complementary form of scholarship.<sup>22,44</sup> Nevertheless, the simplistic hierarchy “systematic review good; narrative review less good” persists in some circles. The under-acknowledged limitations of systematic reviews, along with missed opportunities for undertaking and using narrative reviews to extend understanding within a field, risks legitimising and perpetuating a narrow and unexciting research agenda and contributing to research waste. We call upon policymakers and clinicians (who seek to ensure that their decisions are evidence-based, but who may have been seduced by a spurious hierarchy of secondary evidence) and on research commissioners (whose decisions will shape the generation of the future evidence base) to re-evaluate the low status currently afforded to narrative reviews.

## AUTHORS' CONTRIBUTIONS

TG was invited to submit a paper on a topic of her choice to EJCI by the editor. She suggested this topic to ST and

KM and wrote an initial outline for the paper. All authors then contributed iteratively and equally to the development of ideas and refinement of the paper.

## REFERENCES

- Mulrow CD, Thacker SB, Pugh JA. A proposal for more informative abstracts of review articles. *Ann Intern Med.* 1988;108:613-615.
- Faggion CM, Bakas NP, Wasiak J. A survey of prevalence of narrative and systematic reviews in five major medical journals. *BMC Med Res Methodol.* 2017;17:176.
- Higgins JP, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*, vol. 4. Chichester, UK: John Wiley & Sons; 2011.
- Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol.* 2007;7:10.
- Whiting P, Savović J, Higgins JP, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol.* 2016;69:225-234.
- Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med.* 2009;6:e1000100.
- Pedersen BK. Anti-inflammatory effects of exercise: role in diabetes and cardiovascular disease. *Eur J Clin Invest.* 2017;47:600-611.
- Williams R, Aspinall R, Bellis M, et al. Addressing liver disease in the UK: a blueprint for attaining excellence in health care and reducing premature mortality from lifestyle issues of excess consumption of alcohol, obesity, and viral hepatitis. *Lancet.* 2014;384:1953-1997.
- Cannon JW. Hemorrhagic Shock. *N Eng J Med.* 2018;378:370-379.
- Sackett DL, Rosenberg WM, Gray JM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ.* 1996;312:71-72.
- Alper BS, Haynes RB. EBHC pyramid 5.0 for accessing preappraised evidence and guidance. *BMJ.* 2016;21:123-125.
- Whitty CJ. What makes an academic paper useful for health policy? *BMC Med.* 2015;13:301.
- Malterud K. The art and science of clinical knowledge: evidence beyond measures and numbers. *The Lancet.* 2001;358:397-400.
- Boell SK, Cecez-Kecmanovic D. A hermeneutic approach for conducting literature reviews and literature searches. *CAIS.* 2014;34:12.
- Nevo I, Slonim-Nevo V. The myth of evidence-based practice: towards evidence-informed practice. *Br J Soc Work.* 2011;41:1176-1197.
- Morton S, Berg A, Levit L, Eden J. *Finding What Works in Health Care: Standards for Systematic Reviews*. Washington, DC: National Academies Press; 2011.
- Dixon-Woods M, Cavers D, Agarwal S, et al. Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Med Res Methodol.* 2006;6:35.
- Greenhalgh T. Outside the box: why are cochrane reviews so boring? *Br J Gen Pract.* 2012;62:371-371.
- Ioannidis J. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q.* 2016;94:485-514.
- Egger M, Smith GD, Altman D. *Systematic Reviews in Health Care: Meta-Analysis in Context*. Oxford: Wiley; 2008.
- Lewin S, Glenton C, Oxman AD. Use of qualitative methods alongside randomised controlled trials of complex healthcare interventions: methodological study. *BMJ.* 2009;339:b3496.
- Petticrew M, Rehfuess E, Noyes J, et al. Synthesizing evidence on complex interventions: how meta-analytical, qualitative, and mixed-method approaches can contribute. *J Clin Epidemiol.* 2013;66:1230-1243.
- Hong QN, Pluye P, Bujold M, Wassef M. Convergent and sequential synthesis designs: implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. *Syst Rev.* 2017;6:61.
- Pluye P. Mixed kinds of evidence: synthesis designs and critical appraisal for systematic mixed studies reviews including qualitative, quantitative and mixed methods studies. *Evid-Based Med.* 2015;20:79.
- Pluye P, Hong QN, Bush PL, Vedel I. Opening-up the definition of systematic literature review: the plurality of worldviews, methodologies and methods for reviews and syntheses. *J Clin Epidemiol.* 2016;73:2-5.
- MacLure M. 'Clarity bordering on stupidity': where's the quality in systematic review? *J Edu Policy.* 2005;20:393-416.
- Wong G, Greenhalgh T, Westhorp G, Buckingham J, Pawson R. RAMESES publication standards: realist syntheses. *BMC Med.* 2013;11:21.
- Wong G, Greenhalgh T, Westhorp G, Buckingham J, Pawson R. RAMESES publication standards: meta-narrative reviews. *BMC Med.* 2013;11:20.
- Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O, Peacock R. Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review. *Soc Sci Med.* 2005;61:417-430.
- Paterson BL, Canam C. *Meta-study of Qualitative Health Research: A Practical Guide to Meta-analysis and Meta-synthesis*, vol. 3. London, UK: Sage; 2001.
- Thorne S. Metasynthetic madness: what kind of monster have we created? *Qual Health Res.* 2017;27:3-12.
- Whittemore R, Knaf K. The integrative review: updated methodology. *J Adv Nurs.* 2005;52:546-553.
- Torraco RJ. Writing integrative literature reviews: guidelines and examples. *Hum Resour Dev Rev.* 2005;4:236-367. INSERT
- Saunders MN, Trojon C. On the attributes of a critical literature review. *Coaching.* 2011;4:158-162.
- Randolph JJ. A guide to writing the dissertation literature review. *PARE.* 2009;14:1-13.
- Byrne JA. Improving the peer review of narrative literature reviews. *Res Integr Peer Rev.* 2016;1:12.
- Thorne S, Sawatzky R. Particularizing the general: Sustaining theoretical integrity in the context of an evidence-based practice agenda. *Adv Nurs Sci.* 2014;37:5-18.
- Merton RK. Science and the social order. *Philos Sci.* 1938;5:321-337.
- Webster J, Watson RT. Analyzing the past to prepare for the future: writing a literature review. *MIS Q.* 2002;26:xiii-xxiii.
- Malterud K, Bjelland AK, Elvbakken KT. Evidence-based medicine—an appropriate tool for evidence-based health policy? A case study from Norway. *Health Res Policy Syst.* 2016;14:15.
- Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? *BMJ.* 2014;348:g3725.

42. Levin A. The cochrane collaboration. *Ann Intern Med.* 2001;135:309-312.
43. Greenhalgh T, Shaw S, A'Court C. Understanding heart failure; explaining telehealth—a hermeneutic systematic review. *BMC Cardiovasc Disord.* 2017;17:156.
44. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesising qualitative and quantitative evidence: a review of possible methods. *J Health Serv Res Policy.* 2005;10:45-53.

**How to cite this article:** Greenhalgh T, Thorne S, Malterud K. Time to challenge the spurious hierarchy of systematic over narrative reviews?. *Eur J Clin Invest.* 2018;48:e12931. <https://doi.org/10.1111/eci.12931>