



# What does expert opinion in guidelines mean? a meta-epidemiological study

Oscar J Ponce,<sup>1</sup> Neri Alvarez-Villalobos,<sup>1</sup> Raj Shah,<sup>2</sup> Khaled Mohammed,<sup>3</sup> Rebecca L Morgan,<sup>4</sup> Shahnaz Sultan,<sup>5</sup> Yngve Falck-Ytter,<sup>6</sup> Larry J Prokop,<sup>7</sup> Philipp Dahm,<sup>8</sup> Reem A Mustafa,<sup>9</sup> Mohammad H Murad<sup>1</sup>

10.1136/ebmed-2017-110798

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/ebmed-2017-110798>).

<sup>1</sup>Evidence-based Practice Center, Mayo Clinic, Rochester, Minnesota, USA

<sup>2</sup>Department of Medicine, University of Missouri-Kansas City, Kansas, Missouri, USA

<sup>3</sup>Pediatric Residency Program, University of Minnesota, Minneapolis, Minnesota, USA

<sup>4</sup>Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Canada

<sup>5</sup>Division of Gastroenterology, Hepatology, and Nutrition, University of Minnesota, Center for Chronic Diseases Outcomes Research, Minneapolis Veterans Affairs Healthcare System, Minneapolis, Minnesota, USA

<sup>6</sup>Division of Gastroenterology, Case Western Reserve University, Cleveland, Ohio, USA

<sup>7</sup>Mayo Clinic Libraries, Rochester, Minnesota, USA

<sup>8</sup>Department of Urology, University of Minnesota, Minneapolis, Minnesota, USA

<sup>9</sup>Division of Nephrology and Hypertension, University of Kansas Medical Center, Kansas City, Kansas, USA

Correspondence to:

**Dr Mohammad H Murad**, Evidence-based Practice Center, Mayo Clinic, Rochester, MN 55905, USA; [murad.mohammad@mayo.edu](mailto:murad.mohammad@mayo.edu)

## Abstract

Guidelines often use the term expert opinion (EO) to qualify recommendations. We sought to identify the rationale and evidence type in EO recommendations. We searched multiple databases and websites for contemporary guidelines published in the last decade that used the term EO. We identified 1106 references, of which 69 guidelines were included (2390 recommendations, of which 907 were qualified as EO). A rationale for using EO designation was not provided in most (91%) recommendations. The most commonly cited evidence type was extrapolated from studies that did not answer guideline question (40% from randomised trials, 38% from observational studies and 2% from case reports or series). Evidence extrapolated from populations that were different from those addressed in the guideline was found in 2.5% of EO recommendations. We judged 5.6% of EO recommendations as ones that could have been potentially labelled as good practice statements. None of the EO recommendations were explicitly described as being solely dependent on the clinical experience of the panel. The use of EO as a level of evidence in guidelines remains common. A rationale for such use is not explicitly provided in most instances. Most of the time, evidence labelled as EO was indirect evidence and occasionally was very low-quality evidence derived from case series. We posit that the explicit description of evidence type, as opposed to using the label EO, may add clarity and transparency and may ultimately improve uptake of recommendations.

## Introduction

Clinical practice guidelines are systematically developed statements that intend to assist clinicians and patients in making decisions about appropriate healthcare in specific circumstances.<sup>1,2</sup> Empirical evidence shows that adherence to guidelines improves patient outcomes.<sup>3-5</sup> One of the essential requirements of a trustworthy guideline<sup>2</sup> is that it should be based on systematic reviews of the best available evidence and includes assessment of the quality of evidence. This construct, quality of evidence, has been described as the certainty that a true effect lies on one side of a specified threshold or within a chosen range.<sup>6</sup> The GRADE (Grading of Recommendations, Assessment, Development and Evaluation) Working Group has provided an approach for evaluating the quality of evidence depending on several discrete domains, which include study limitations, indirectness, imprecision, inconsistency and publication bias.<sup>5</sup>

## Summary box

### What is already known about this subject?

- Many guidelines use the term expert opinion in recommendations when evidence is considered to be insufficient.
- Guidelines that use the term expert opinion are often viewed as less rigorous.
- It is unclear what expert opinion means.

### What are the new findings?

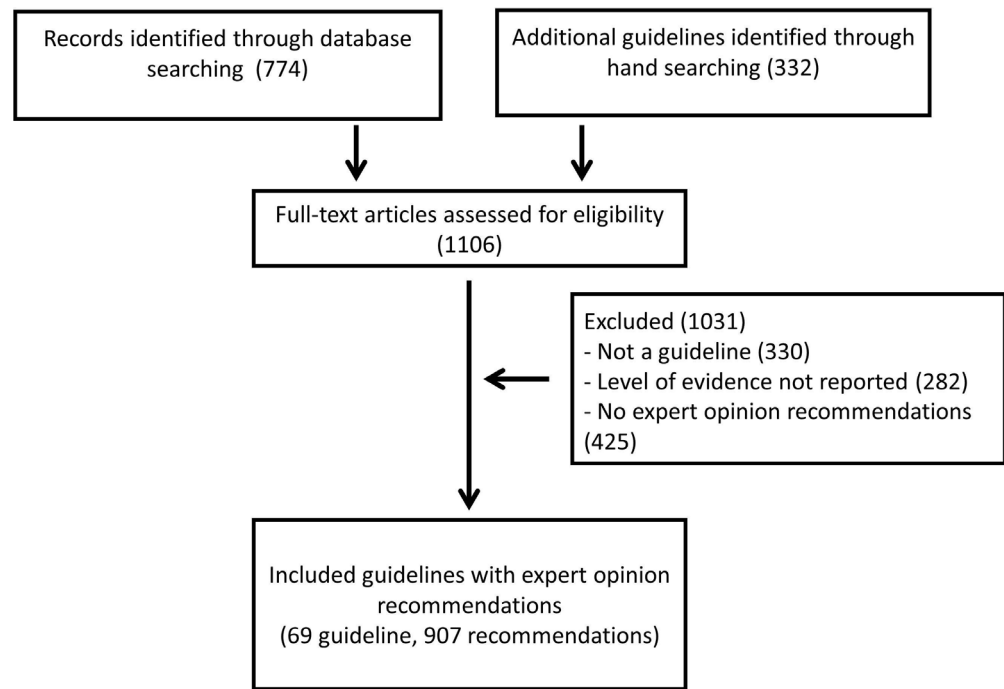
- Many guidelines that used the term expert opinion did not provide a rationale.
- Expert opinion recommendations often contained evidence of various types.
- Expert opinion recommendations were most often used to describe indirect evidence.

### How might it impact clinical practice in the foreseeable future?

- Explicit description of evidence type, such as indirect evidence, may add clarity and transparency and may ultimately improve uptake of recommendations.

Nevertheless, in the daily language of clinicians and in many scientific publications including guidelines, the term expert opinion (EO) is commonly used. Intuitively, clinical decision making requires expertise, which makes depending on EO logical. Several depictions of the evidence pyramid consider EO as a level of evidence and place it at the bottom of the pyramid as a unique category, or combined with preclinical studies and case reports, implying low validity. However, the term 'opinion' is defined in dictionaries as a 'generally held view', a 'belief' or a 'judgement formed in the mind about a particular matter'.<sup>7</sup> These definitions are not fully congruent with the definition of evidence as an empirical observation. Clinical practice guidelines that are based on EO may be viewed negatively and can be considered less trustworthy.<sup>5</sup> Yet clinical expertise remains critical in decision making. The GRADE approach to EO has therefore been that EO should not be used as a separate category of evidence; rather, clinical expertise is considered as an essential ingredient for interpreting evidence of all study types and to formulate evidence-based recommendations.

Therefore, we sought to critically evaluate contemporary clinical practice guidelines in which the supporting evidence for specific recommendations was described as



**Figure 1** Guideline selection process.

EO to determine the rationale of guideline developers for using EO and explore factors that are associated with EO use. We hypothesised that EO may reflect several types of evidence that can potentially be described using better semantics and rated more appropriately using the standard quality of evidence domains. More accurate grading of evidence may lead to more clarity and transparency and may reduce the perceived tension between evidence and expertise.

### Methodology

This meta-epidemiological study uses a systematic review approach and follows appropriate reporting guidelines.<sup>8</sup>

### Search strategy

A comprehensive search developed by an expert librarian (LJP) with input from study investigators was conducted from 1 January 2010 to 3 June 2016 to capture contemporary published guidelines. We searched the following databases: Ovid MEDLINE In-Process & Other Non-Indexed Citations, Ovid MEDLINE and Ovid EMBASE. Controlled vocabulary supplemented with keywords was used to search for clinical practice guidelines (online supplementary appendix). We also conducted hand searches (internet search) for websites of several professional societies and entities that we knew produce guidelines and use EO designation (Infectious Diseases Society of America, Centers for Disease Control and Prevention, American Psychological Association, American Academy of Family Physicians, American Society for Reproductive Medicine, American Society of Clinical Oncology, American Urological Association, American Heart Association and American Diabetes Association).

Guidelines were included if they reported EO as part of their quality of evidence assessment system or if the

strength of recommendation grading system had a level in which they state that EO was the only basis of the recommendation. We excluded guidelines that did not use EO as a grade. We piloted the process of study selection using five guidelines to ascertain common understanding among reviewers.

### Data collection

Teams of paired reviewers independently extracted data using a prepiloted screening and abstraction form. We assigned several a priori categories for the type of evidence cited as EO, which included extrapolation from research studies (randomised, observational, case series/case reports or basic science), and for providing the rationale for using EO recommendations. We also assessed whether EO recommendations were better described as good practice statements (ie, statements supported by a wealth of indirect evidence and no direct evidence<sup>9</sup>). We estimated the observed agreement among reviewers for assessing the type of evidence cited in EO recommendations. Consensus was reached by discussion with a third reviewer. We extracted the following variables: publication year, society/organisation that developed the guideline, guideline field, presence of methodology experts in the development of the guideline (identified through authorship contribution, affiliation or by searching authors' background on Google), declared conflicts of interest, definition of EO, use of systematic reviews, and characteristics of EO recommendations such as the stated rationale for using EO and the type of evidence cited as EO.

### Statistical analysis

We used descriptive statistics to summarise the epidemiology of EO recommendations (prevalence, trend over time) and the rationale and evidence type cited in EO

| Table 1 Characteristics of the included guidelines         |                |      |
|--|----------------|------|
| Characteristics  | Guidelines (n) | %    |
| <b>Year</b>  |                |      |
| 2010   | 6              | 8.7  |
| 2011   | 7              | 10.1 |
| 2012   | 7              | 10.1 |
| 2013   | 7              | 10.1 |
| 2014   | 9              | 13.0 |
| 2015   | 16             | 23.2 |
| 2016   | 17             | 24.6 |
| <b>Methodologist included in guideline</b>                 |                |      |
| Yes  | 24             | 34.8 |
| No   | 9              | 13.0 |
| Unclear  | 36             | 52.2 |
| <b>Declared financial conflict of interest</b>             |                |      |
| Yes  | 63             | 91.3 |
| No   | 2              | 2.9  |
| Unclear  | 4              | 5.8  |
| <b>Systematic review use</b>                               |                |      |
| Guideline cited published reviews                          | 23             | 33.3 |
| Review embedded in guideline                               | 21             | 30.4 |
| Guideline had a rapid review or an updated literature scan | 18             | 26.1 |
| Guideline had reviews commissioned                         | 7              | 10.1 |
| <b>Area of clinical specialty</b>                          |                |      |
| Endocrinology  | 40             | 58.0 |
| Urology  | 13             | 18.8 |
| Orthopaedic surgery  | 5              | 7.2  |
| Cardiology   | 3              | 4.3  |
| Other*   | 8              | 11.6 |

\*Gynaecology, neonatology, pulmonology, psychiatry, rheumatology, surgery and otorhinolaryngology.

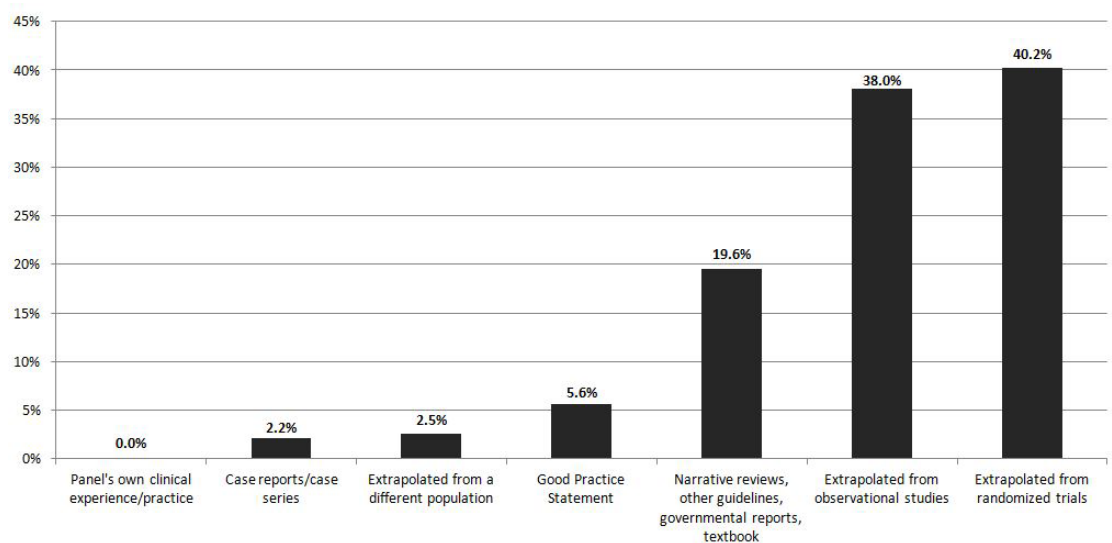
recommendations. We explored characteristics of EO recommendations associated with citing evidence (rate of EO per guideline, rationale for EO, presence of a methodologist on the guideline team and type of systematic review process used in the guideline). We used Pearson's  $\chi^2$  test or Fisher's exact test to compare nominal variables. For continuous variables, we used one sample Kolmogorov-Smirnov test to test the normality of distribution. Unpaired Student's or Mann-Whitney U test was used to compare between groups. A two-tailed p value <0.05 was considered statistically significant. The statistical analyses were done using SPSS V.20.0.

## Results

### Search results

We assessed 1106 references for eligibility. We included 69 guidelines that provided 2390 recommendations, of which 907 (37.9%) were labelled as having a level of evidence designated as EO. The process of guideline selection is depicted in figure 1.

The proportion of guidelines with EO recommendations increased from 2010 to 2016, from 6 (8.7%) to 17 (24.6%), respectively. Twenty-four (34.8%) out of 69 included a methodologist in the development of the guideline, and almost all the guidelines declared to have at least one author with a financial conflict of interest (91.3%). All guidelines used systematic reviews to select evidence (59.4% used published reviews or conducted rapid reviews and 40.6% commissioned or conducted their own). The most common clinical fields addressed by guidelines were endocrinology (58%), urology (18.8%) and orthopaedic surgery (7.2%). The characteristics of these guidelines are summarised in table 1. The observed agreement among reviewers about the



**Figure 2** Evidence type cited as expert opinion. Categories overlap and proportions do not sum to 100%. The denominator of 'Good practice statement' and 'Extrapolated from a different population' is all expert opinion recommendations, whereas other categories have the denominator of expert opinion recommendations that reported evidence type.

**Table 2** The effect of a methodologist role and systematic review process on expert opinion recommendations

|   | Methodologist   |   | Total       | p      |
|---|---|---|-------------|--------|
|   | Yes (n=24)  | No/Unclear (n=45)   |             |        |
| Mean rate of EO recommendations per guideline* (SD)   | 0.30 (0.18)   | 0.49 (0.33)   | 0.43 (0.30) | 0.03   |
| Mean rate of EO recommendations that cited evidencet (SD)                                       | 0.51 (0.31)   | 0.22 (0.2)  | 0.32 (0.3)  | <0.001 |
| Number of guidelines that used 'Lack of evidence' as a rationale for the EO recommendations (%) | 14 (50)   | 14 (31.1)   | 28 (40.6)   | 0.04   |
| Number of CPG that had its own SR in it or had an SR commissioned (%)                           | 21 (87.5)   | 7 (15.6)  | 28 (40.6)   | <0.001 |
| Number of CPG that cited published SRs, had a rapid review or an updated literature scan (%)    | 3 (12.5)  | 38 (84.4)   | 41 (59.4)   |        |
|   | CPG that had its own SR in it or had SR commissioned (n=28) | CPG that cited published SRs or had a rapid review or an updated literature scan (n=41) | Total       | p      |
| Mean rate EO recommendations that cited evidencet (SD)  | 0.54 (0.31)   | 0.18 (0.19)   | 0.32 (0.30) | <0.001 |
|   | Declaring conflicts of interest                             |   | Total       | p      |
|   | Yes (n=63)  | No/Unclear (n=6)  |             |        |
| Mean rate of EO recommendations that cited evidencet (SD)                                       | 0.30 (0.29)   | 0.54 (0.36)   | 0.32 (0.3)  | 0.07   |

\*Number of EO recommendations divided by the total number of recommendations per guideline.

tNumber of EO recommendations that cited evidence divided by the total number of EO recommendations per guideline.

CPG, clinical practice guideline; EO, expert opinion; SR, systematic review.

underlying type of evidence used in EO recommendations was 77%.

### Rationale and evidence type labelled as EO

Only seven societies provided a definition of EO (online supplementary table 1). These definitions suggested lack of evidence, evidence development in the context of research, depending on physiology, bench research or panel's own clinical experience. The majority of EO recommendations did not provide a rationale for using EO (828/907, 91%) and only a few (79/907, 8.7%) stated lack of evidence as rationale.

Evaluation of 30.4% (276/907) EO recommendations in which any type of evidence was cited reveals that the most common reason for using EO was extrapolation of evidence from studies that did not quite answer the question of the guideline (40.2% from randomised trials, 38% from observational studies). Other less common reasons were deriving inferences from case reports and case series (2.2%) (figure 2). A few of EO recommendations (2.5%) stated that their evidence was extrapolated from a different population, and we judged 5.6% of EO recommendations as ones that could have potentially been described as good practice statements. In online supplementary table 2, we provide a list of all included guidelines.

Having a methodologist as a part of a guideline team was associated with increased rate of EO recommendations per guideline (p=0.03), with using 'lack of evidence' as a rationale for the EO recommendations (p=0.04), with increased rate of EO recommendations per guideline that cited evidence (p=0.01) and with guidelines conducting or commissioning their own systematic reviews (p<0.01). Guidelines that commissioned or conducted their own

systematic reviews also had increased rate of EO recommendations per guideline that cited evidence (compared with guidelines that did not use systematic reviews or used existing or rapid reviews; p<0.01). Citing evidence in EO recommendations and the proportion of EO per guideline recommendations were not significantly associated with declaring conflicts of interest (p>0.05) (table 2).

## Discussion

### Main findings

We conducted a meta-epidemiological study to evaluate published contemporary clinical practice guidelines in terms of their use of EO as a level of evidence. We found that in this sample of guidelines, EO use was very common. We also found that the majority of EO-based recommendations did not explicitly provide a rationale for the use of EO. In the overwhelming majority of situations where EO recommendations had evidence cited, it seemed that EO was in fact not an opinion, but rather an extrapolation of evidence from other populations or settings. As such, it was not congruent with the definition of opinion and could have been better described, in our view, as indirect evidence.<sup>10</sup> Indirect evidence, which is evidence that does not directly answer our question, may warrant lower certainty, but not always. Evidence can be indirect in several ways such as when patients in the study differ from those of interest, when the intervention tested in the study differs from the intervention of interest, or when the study uses a surrogate endpoint. Decisions regarding the importance of indirectness (how much it lowers the quality of evidence) depend on our understanding of whether biological and other factors

are sufficiently different to the extent that one might expect substantial differences in the direction or magnitude of effect.<sup>10</sup>

Some of the EO recommendations were actually dependent on case series and case reports, which usually (but not always) warrants very low certainty in the evidence. We also encountered EO recommendations that could have been described as good practice statements (statements of action supported by a wealth of indirect evidence and little or no direct evidence, considered by clinicians as non-controversial and do not require evidence summaries).<sup>9</sup> This suggests that guideline developers are often conflating EO and good practice statements. We were surprised to not find EO recommendations in which guideline developers relied on their own clinical practice and experience (ie, based the recommendation on their unsystematic observations in practice). However, since many EO recommendations did not provide cited evidence, it is plausible that many of these EOs were based on clinical observations of the experts in their own practice.

This study has also demonstrated that having a methodologist as a part of a guideline team was associated with increased rate of EO recommendations per guideline. This may signify that rating of the quality of evidence was either controversial, sparse or of low quality, which required increased use of experts' input. The presence of a methodologist also increased the likelihood of EO recommendations citing evidence, which is expected and likely relates to the methodologist encouraging experts to be more transparent. We were unable to demonstrate a significant impact of financial conflicts on the use of EO; however, prior literature has demonstrated that experts' opinion is affected by experts' financial ties to industry.<sup>11</sup>

### Implications

This study has demonstrated that most of the time EO is not an opinion; rather, it is indirect evidence or very low-quality evidence. We propose using the correct characterisation and using the well-established categories of quality of evidence (ie, high, moderate, low or very low).<sup>12</sup> Such use is more accurate and transparent for decision makers and users of guidelines. We hypothesise that proper and explicit description of the quality of evidence may improve uptake of recommendations because guideline users may view EO recommendations as less rigorous or less compelling.

While some clinical experts demand to include EO as a source of evidence during guideline development,<sup>13</sup> they define such opinion as:

*'...the opinions of experts are based not only on their personal clinical experiences, but also on their accumulated knowledge from a wide range of sources. These include the expert's personal assessment of the validity of published reports, new knowledge learned at meetings and symposia, awareness of unpublished studies with "negative" results, and knowledge of the (often unreported) practice styles of colleagues in their field of expertise'.<sup>13</sup>*

This definition clearly includes many types of evidence, as opposed to opinion. We propose that each type

should be explicitly stated and appraised using clear labels and to avoid the vague terminology of opinion. This approach is congruent with the GRADE approach. GRADE specifically acknowledges that expertise is required for interpretation of any form of evidence but considers EO to be interpretation of evidence, but not a form of evidence in itself.<sup>12</sup>

When there is a lack of published evidence, one could aim to systematically summarise observations of experts and trends in their practice. This could be practically achieved by surveying the clinical experts, not about their opinion of what should be recommended, but rather about their observations of patients' outcomes over the years. This approach has been used by some guidelines and is well received by the clinical community when there is no published evidence addressing the questions that underpin some recommendations.<sup>14 15</sup>

The quality of evidence supporting the majority of recommendations is usually low (compared with the number and diversity of guidance that clinicians seek). This low quality, along with the fact that added rigour can lead to strict inclusion criteria of evidence, can both result in a guideline that is out of context or lacking implementation details. Therefore, for a guideline to be useful, clinical experts are needed to contextualise evidence, extrapolate evidence from indirect sources and interpret low-quality evidence. In addition to using the correct label for evidence (ie, indirect evidence instead of EO), guidelines should provide an explicit explanation for how the experts were selected, how they reached consensus and what problems experts were asked to address (eg, to weigh in on strength and limitations of evidence, to contextualise evidence and recommendations, to assess and further inform the guideline through indirect evidence, to offer opinions based on their personal clinical experiences, or to enrich and support the guideline text with good practice statements).

### Limitations and strengths

This study depended on judgements made by our investigators when assigning a rationale retrospectively to published text. In addition, appraising the quality of a study or the quality of evidence involves making a judgement.<sup>16 17</sup> The rationale for EO was not provided in a large number of recommendations, limiting conclusions to guidelines that have provided such rationale. In some recommendations it was not clear whether the EO was a designation for the level of evidence or the strength of recommendations. This differentiation, which is a critical component of the GRADE system, was less clear in guidelines using other guideline-development approaches. Another limitation is that a large proportion of EO recommendations were developed by a single professional society. Lastly, the definition of good practice statement as a distinct type of recommendations is fairly recent and some guideline developers may not be aware of such designation.<sup>18</sup>

The strengths of this meta-epidemiological study relate to the comprehensive search and rigorous approach of reviewing studies by independent pairs of reviewers. To our knowledge, this is the first study that attempts to define EO used in guidelines.

## Conclusion

Clinical experts are essential to provide context to the evidence and aid in its interpretation. Experts' engagement in guideline development is particularly critical when evidence is being extrapolated from indirect sources. We suggest avoiding the use of the term expert opinion and replacing it with the appropriate standard terminology for describing the quality of evidence. This approach may improve the uptake of recommendations by improving clarity and transparency.

**Contributors** MHM conceived the study. OJP, NA-V, RS and KM selected studies and extracted data. RLM, SS, YF-Y, LJP, PD and RAM critically revised the manuscript and helped in interpretation. OJP and NA-V conducted analysis. MHM is the grantor of the work.

**Competing interests** None declared.

**Provenance and peer review** Not commissioned; externally peer reviewed.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

## References

- Field M, Lohr K. *Clinical practice guidelines: directions for a new program*, Institute of Medicine. Washington, DC: National Academy Press, 1990.
- Institute of Medicine. *Clinical practice guidelines we can trust*. Washington DC: National Academies Press, 2011.
- Baltesca MA, LaGuardia JC, Lee PC, et al. An electronic order set for acute myocardial infarction is associated with improved patient outcomes through better adherence to clinical practice guidelines. *J Hosp Med* 2014;9:155–61.
- Oh SW, Lee HJ, Chin HJ, et al. Adherence to clinical practice guidelines and outcomes in diabetic patients. *Int J Qual Health Care* 2011;23:413–9.
- Murad MH. Clinical practice guidelines: a primer on development and dissemination. *Mayo Clin Proc* 2017;92:423–33.
- Hulterantz M, Rind D, Akl EA, et al. The GRADE Working Group clarifies the construct of certainty of evidence. *J Clin Epidemiol* 2017;87:4–13.
- Merriam-Webster. Definition of opinion. <https://www.merriam-webster.com/dictionary/opinion> (accessed 23 Jun 2017).
- Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. *Evid Based Med* 2017;22:139–42.
- Guyatt GH, Alonso-Coello P, Schünemann HJ, et al. Guideline panels should seldom make good practice statements: guidance from the GRADE Working Group. *J Clin Epidemiol* 2016;80:3–7.
- Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. rating the quality of evidence--indirectness. *J Clin Epidemiol* 2011;64:1303–10.
- Wang AT, McCoy CP, Murad MH, et al. Association between industry affiliation and position on cardiovascular risk with rosiglitazone: cross sectional systematic review. *BMJ* 2010;340:c1344.
- Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. rating the quality of evidence. *J Clin Epidemiol* 2011;64:401–6.
- Eibling D, Fried M, Blitzer A, et al. Commentary on the role of expert opinion in developing evidence-based guidelines. *Laryngoscope* 2014;124:355–7.
- Mustafa RA, Zimmerman D, Rioux JP, et al. Vascular access for intensive maintenance hemodialysis: a systematic review for a Canadian Society of Nephrology clinical practice guideline. *Am J Kidney Dis* 2013;62:112–31.
- Nesrallah GE, Mustafa RA, MacRae J, et al. Canadian Society of Nephrology guidelines for the management of patients with ESRD treated with intensive hemodialysis. *Am J Kidney Dis* 2013;62:187–98.
- Ip S, Kitsios GD, Chung M, et al. *A process for robust and transparent rating of study quality: phase 1*. Rockville(MD), 2011.
- Murad MH, Mustafa R, Morgan R, et al. Rating the quality of evidence is by necessity a matter of judgment. *J Clin Epidemiol* 2016;74:237–8.
- Guyatt GH, Schünemann HJ, Djulbegovic B, et al. Guideline panels should not GRADE good practice statements. *J Clin Epidemiol* 2015;68:597–600.