

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Do guidelines offer implementation advice to target users?: A systematic review of guideline applicability.
AUTHORS	Gagliardi, Anna; Brouwers, Melissa

VERSION 1 - REVIEW

REVIEWER	Susanne Rabady Paracelsus Medical University, Salzburg, Austria
REVIEW RETURNED	07-Dec-2014

GENERAL COMMENTS	Revision satisfactory.
-------------------------	------------------------

REVIEWER	Sabine van der Veer European Renal Best Practice (Ghent, Belgium)
REVIEW RETURNED	18-Dec-2014

GENERAL COMMENTS	<p>Gagliardi and Brouwers resubmitted a revised version of their paper on the applicability of guidelines published after 2008. They addressed some but not all of my concerns satisfactorily. Especially the analysis strategy still lacks rigor to adequately answer the research questions, even though the data to perform more robust analyses are readily available.</p> <p>1. In their response letter, the authors stated that they feel studying individual guidelines was outside the scope of their study, but this argument is far from being convincing. Why analyse collated data if the underlying raw data are available? I quickly scanned the 12 included reviews, and most (if not all) seem to provide the information that Gagliardi and Brouwer defined as relevant to their research questions (AGREE domain scores, year of publication, country, and type of guideline developer) at the individual guideline level; the authors of the reviews could be contacted in case of missing data. Using the guideline-level raw data would allow answering the research questions much more accurately, because the 76 guidelines published before 2008 could be excluded from all analysis. More specifically:</p> <p>a. It would enable recalculation of review-level and overall summary statistics in Table 2, and thus drawing of more valid conclusions on the state of guideline applicability since 2008. In the current version, the results and conclusions are diluted due to 30% of guidelines published in an era that is out of scope for the current study; this is especially problematic since the authors expect year of publication to affect applicability. The authors stated this as a limitation in the Discussion (pg.12, line 44), but I feel that this is insufficient knowing that the data to address the limitation are readily available.</p>
-------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

	<p>b. For research question 2 the authors refrain from doing any statistics due to 'the small number of eligible studies' and instead 'observe trends between applicability and' (pg.9, line 15-25). This methodological choice leads to a disappointing section on 'Factors influencing applicability' (pg. 11) that does not provide an answer. Doing a multilevel linear regression analysis using the available guideline-level data (n=178; 'systematic review' would be one of the levels in the model) would allow exploring associations between the applicability domain score and the factors as postulated by the authors.</p> <p>c. The statement on pg. 12 (line 49) "...and if we had imposed such strict eligibility criteria..." was posed as a reason not to do a guideline-level analysis. It seems to refer to excluding reviews in case they show any overlap with Alonso-Coello et al. or Knai et al. Assuming my interpretation is correct, I do not consider this a valid criterion since this overlap is limited: 3 guidelines in Knai and none in Alonso were published in 2008 or later. It might therefore be sufficient to report how many guidelines included in the current meta-review were also included in Knai's study. In case I misinterpreted, the statement needs additional explanation.</p> <p>2. Also, some of my concerns regarding the inclusion criteria remain. Exclusion of reviews that did not report the applicability domain score is obvious because they would not contribute to answering any of the research questions. However:</p> <p>a. It is still unclear whether reviews that did report domain scores but 'primarily compared underlying evidence, development methods or recommendations across guidelines' were excluded. If that is the case, the authors should better explain this choice in terms of expected distortion of findings. The same is true for excluding studies that served as a 'baseline needs assessment in countries new to guideline development'.</p> <p>b. Why not include reviews of 'guidelines sampled from one organisation'? Adding 'organisation' as a level in the regression analysis (see point 1b) would account for the fact that guidelines produced by one organisation are more likely to have similar applicability scores than guidelines produced by different organisations.</p> <p>c. I appreciate that the authors made a conscious decision to exclude 2 reviews that reported applicability domain scores but not all of the other domain scores. Yet, looking at the research questions, they did not convince me that this a good choice. These 2 studies will contain multiple guidelines that could be included in the analysis, adding robustness to the findings. They would have 'missing data' for analysing the association between domains, but this is only part of research question 2. All the other associations under investigation in question 2 + entire question 1 would still be addressed.</p> <p>Other points:</p> <p>3. The argument around why developers should invest in increasing the applicability of their guidelines (as defined by AGREE) is very focused on implementation tools, which is only one item in the applicability domain.</p> <p>a. In contrast to ref 9/10, ref 11 does not seem appropriate in this context, because CDS systems are unlikely to be taken into account in AGREE appraisals.</p> <p>b. It would help if the authors could add some studies showing the potential benefit of providing information/tools as suggested for the other applicability items.</p>
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

	<p>4. The secondary purpose (pg.7, lines 15-18) should be reformulated by taking out 'may have influenced'. This seems to suggest causality between 'factors' and 'applicability', while the observational nature of the data will only allow identification of associations.</p> <p>5. I agree that the AGREE manual clearly states that appraisers must examine all related documents, and it helps that the authors emphasized this in the introduction. However, the fact that – compared to other AGREE domains—the information to correctly appraise applicability items is more likely to be found outside the guideline document, decreases the reliability of applicability scores. The authors should consider stating this as a limitation of using AGREE scores to assess guideline applicability.</p> <p>6. References in Table 2 require renumbering.</p>
--	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

VERSION 1 – AUTHOR RESPONSE

REVIEWER #1 - Susanne Rabady
Revision satisfactory (no edits required)

REVIEWER #2 - Sabine van der Veer

Point #1

To address these legitimate concerns we revised eligibility criteria to include all reviews published in 2008 or later that included any guidelines published in 2008 or later for which domain scores were provided or could be calculated based on data in the manuscript (given time limitation for responding to these comments we did not approach authors to request data), extracted data for guidelines published in 2008 or later, and conducted a multi-level regression analysis to more rigorously identify any associations between applicability score and hypothesized influencing factors. As a result, more studies were eligible, new data tables were included for studies and guideline-level summary statistics and regression analysis, the Results were entirely rewritten, and points in the Discussion, Abstract and Strengths and Limitations statement were edited.

Point #2

We elaborated on these points to justify our decisions, and edits appear in bold type in Methods, Searching and screening section, page 8.

Of the two studies in which only applicability domain scores were examined and reported, one study did not include any guidelines published in 2008 or later and was ineligible, and the other included only two guidelines published in 2008 or later so we made a conscious decision not to include it since we were able to include 8 additional reviews with many more guidelines published in 2008 or later and also reporting all domain scores.

Point #3

We included reference 9 to 11 specifically because they reflect the types of decision-making tools recommended by Greenhalgh, Pronovost and others, and for which there is evidence of impact on guideline use, hence, our interest in guideline implementation tools. There is less evidence available on the association between inclusion of resource considerations and audit criteria in guidelines and guidelines use, however, that is an issue with AGREE, and not with the justification for our study.

Point #4

As suggested, in Introduction, second last sentence on pg 7, we replaced "...to identify factors that may have influenced applicability." with "...to identify factors associated with applicability."

Point #5

In our experience, information reflecting domains other than applicability can also be found outside the guideline, for example, additional documents that describe how the guideline was developed and by who, thus addressing stakeholder involvement, rigour of development and editorial independence. Conversely, many guidelines offer implementation advice as sections or chapters within the guideline document. However, we did add this as a potential limitation to that section of the Discussion on page 14 to 15.

Point #6

References in Table 2 were re-numbered, in particular because we added additional studies.

VERSION 2 – REVIEW

REVIEWER	Sabine van der Veer European Renal Best Practice (Ghent, Belgium)
REVIEW RETURNED	27-Jan-2015

GENERAL COMMENTS	The authors have sufficiently addressed my concerns.
-------------------------	------------------------------------------------------