Evaluation of AHRQ Evidence-based Practice Center (EPC) Program Reports

Haley Holmer, PhD MPH
Scientific Resource Center
AHRQ EPC Program

National Academies Webinar
Updating Standards for Systematic Reviews
February 27, 2026

Disclosures

I have nothing to disclose.

The author is responsible for the content of this presentation. It does not necessarily represent the views of the Agency for Healthcare Research and Quality (AHRQ). No statement in this presentation should be construed as an official position of AHRQ or of the US Department of Health and Human Services. This work was funded by AHRQ through the Scientific Resource Center (Contract No. HHSA 75Q80122C00002).

Outline

1. Overview of the AHRQ Evidence-based Practice Center (EPC) Program
    - Evolution of methods development

2. Evaluation of EPC Program reports

3. Implications for updating the 2011 IOM standards

Agency for Healthcare Research and Quality

Evidence-based Practice Center (EPC) Program

- Established in 1997
- 38 federal partners; 66 private sector partners
- More than 800 reports produced on medications, devices, healthcare services
    - 200+ were used for clinical practice guidelines
    - 35 were used for national coverage determinations

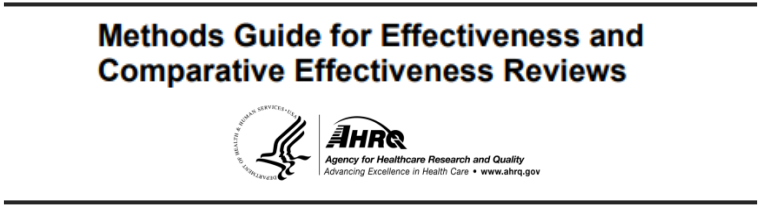
https://effectivehealthcare.ahrq.gov/about/epc-25-years

Agency for Healthcare Research and Quality

Evidence-based Practice Center (EPC) Program

In July 2025, 5-year contracts were awarded to 11 EPCs.

The EPC Program has published 2 influential methods guides.



The first, published in 2001, described methods work for the USPSTF. It was expanded and republished in 2020.

The second, published in 2008 and later published as a series in JCE in 2010, addressed comparative effectiveness reviews of medical interventions.

Decades of methods development in the EPC Program

- Observational and nonrandomized studies (1998-)
- Unpublished literature and reporting bias (2005-)
- Rapid reviews, living reviews (2015-)
- Complex interventions (2015-2016)
- End-user / usability (2017)
- Communication / dissemination of findings (2018-2020)

Evaluation of EPC Program Reports (2018-2023)

Taking stock of EPC Report Quality

Motivation for evaluation:

- Improve the content and readability of EPC reports
- Prioritize methods areas that might need to be addressed
    - Change existing guidance
    - Develop new guidance
    - Conduct trainings with EPCs

Background:

- EPC Program routinely engages in quality improvement
- Editorial review process is lengthy, often with multiple rounds of review
    - Internal review is intended to be a methodologic review
    - Authors have an opportunity to respond to comments before going to peer review

We needed a better understanding of the content/impact of reviewer comments

Approach

We selected 25 EPC reports accepted between 6/5/2018 and 3/2/2023

Excluded:

- Methods reports, technology assessments

Then we abstracted all internal review and external review comments into a database

And applied this idea:

Assess the concordance between internal and external peer review comments

Internal Review	External Review: Methods concerns present	External Review: Methods concerns absent
Methods concerns present	**a** (internal review concerns persist)	**b**
Methods concerns absent	**c** (concerns raised in external review)	**d**

Typology of Problems

"Systematic Reviewlution" Framework

Journal of Clinical Epidemiology

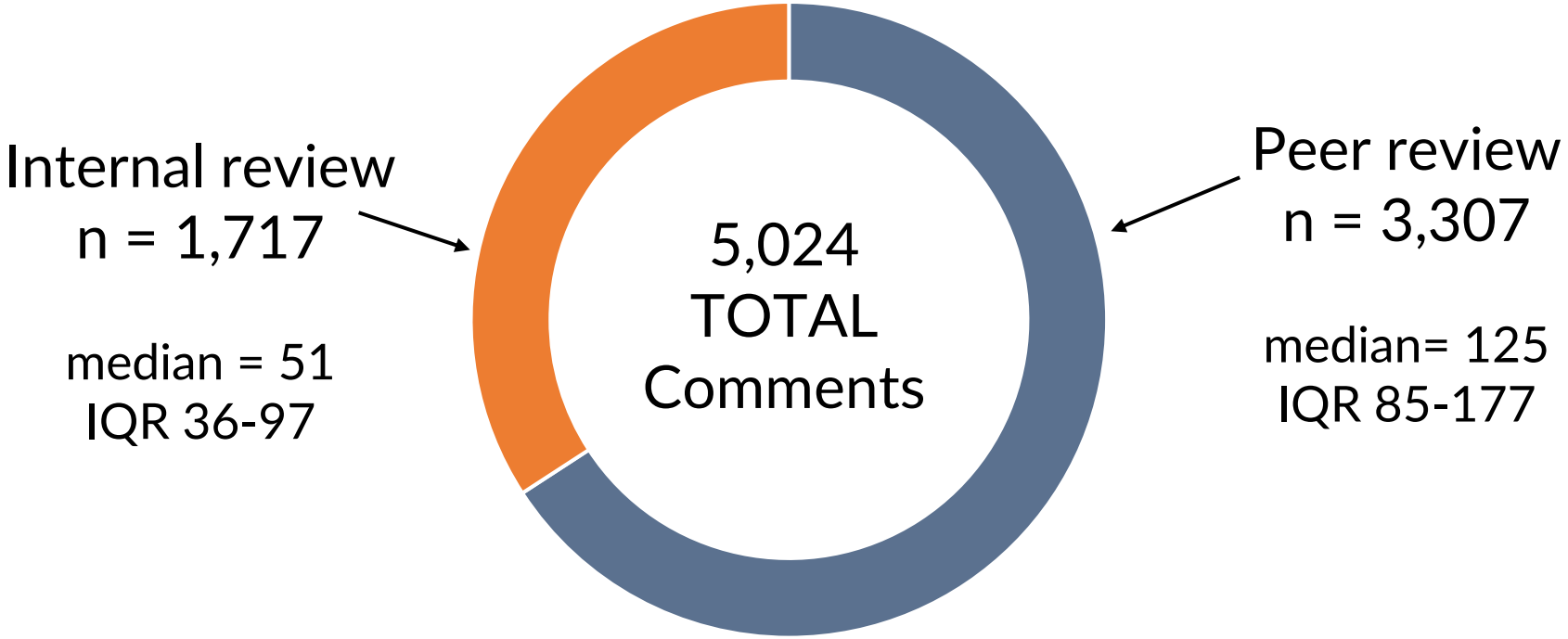
REVIEW ARTICLE

The problems with systematic reviews: a living systematic review

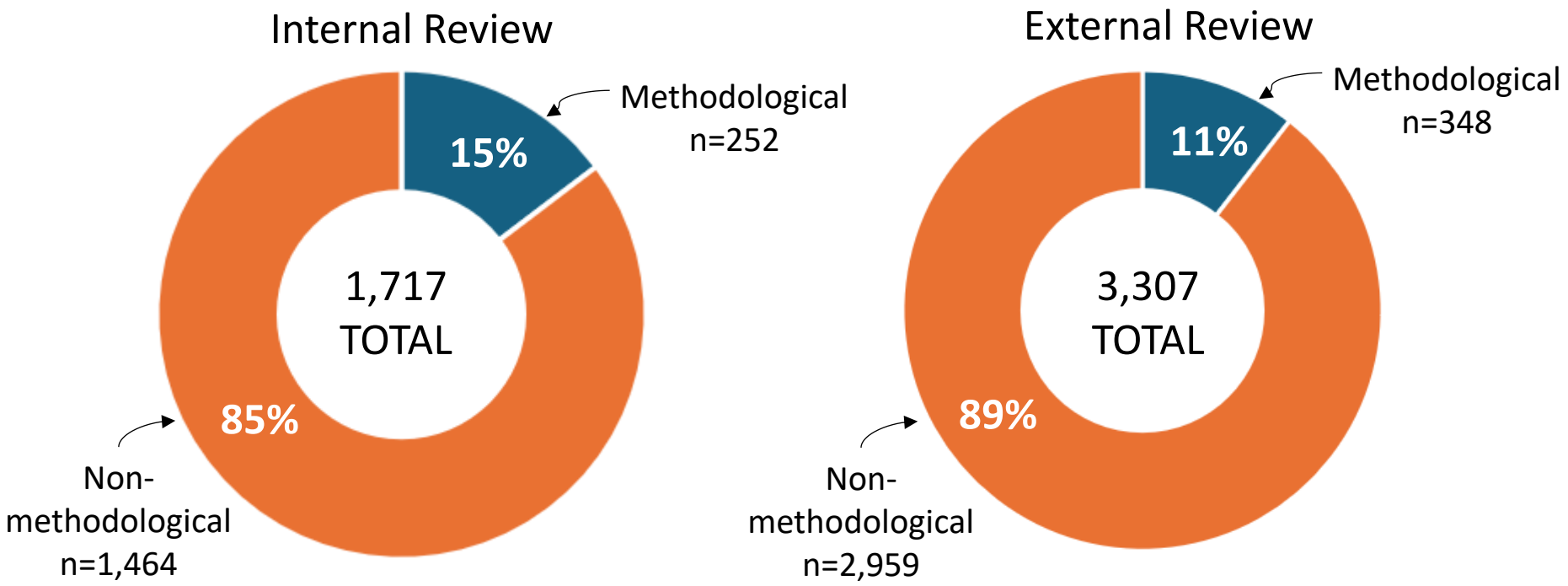
Lesley Uttley[a,*], Daniel S. Quintana[b,c,d], Paul Montgomery[e], Christopher Carroll[a], Matthew J. Page[f], Louise Falzon[a], Anthea Sutton[a], David Moher[g]

- 67 problems included in the framework
- We used it to classify methods concerns in EPC reports

Results

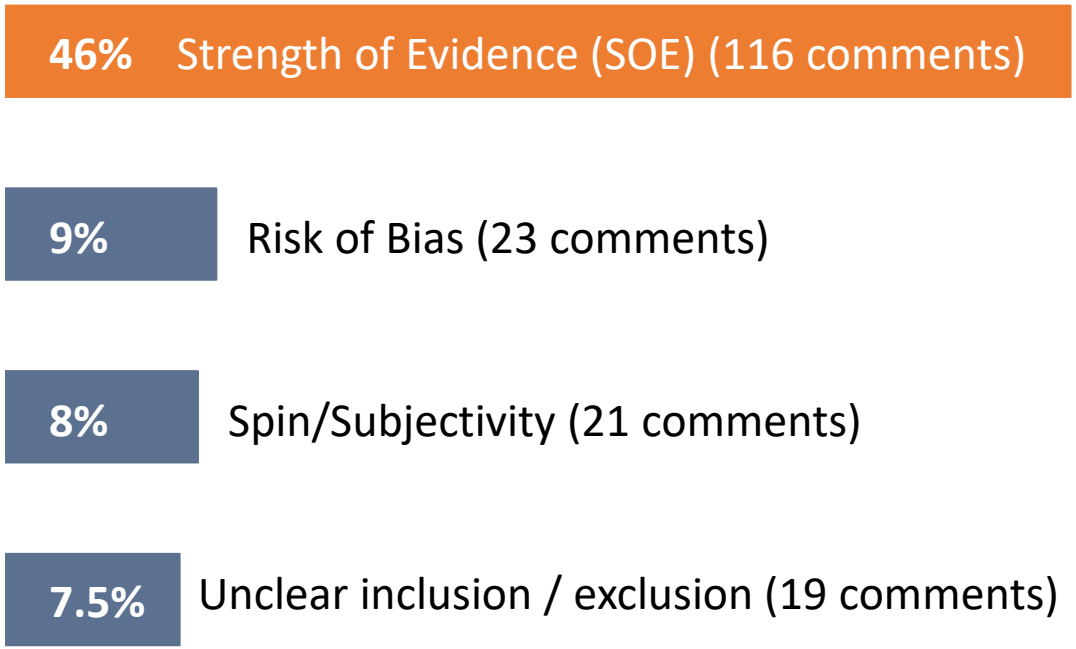

Internal review
n = 1,717
median = 51
IQR 36-97
5,024
TOTAL
Comments
Peer review
n = 3,307
median= 125
IQR 85-177

Most reviewer comments did not address methods.



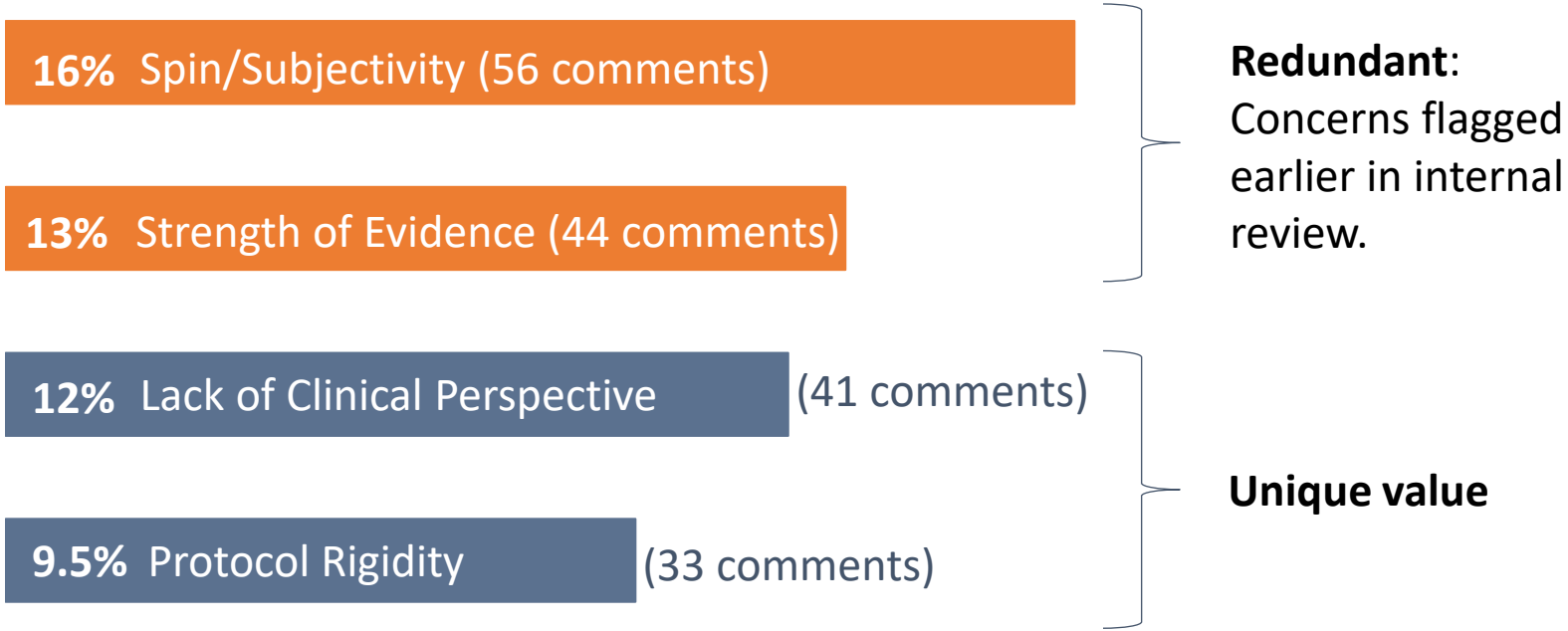
We focused on the comments that had methodological implications.

Internal reviewers focused on evidence grading, risk of bias, and spin.

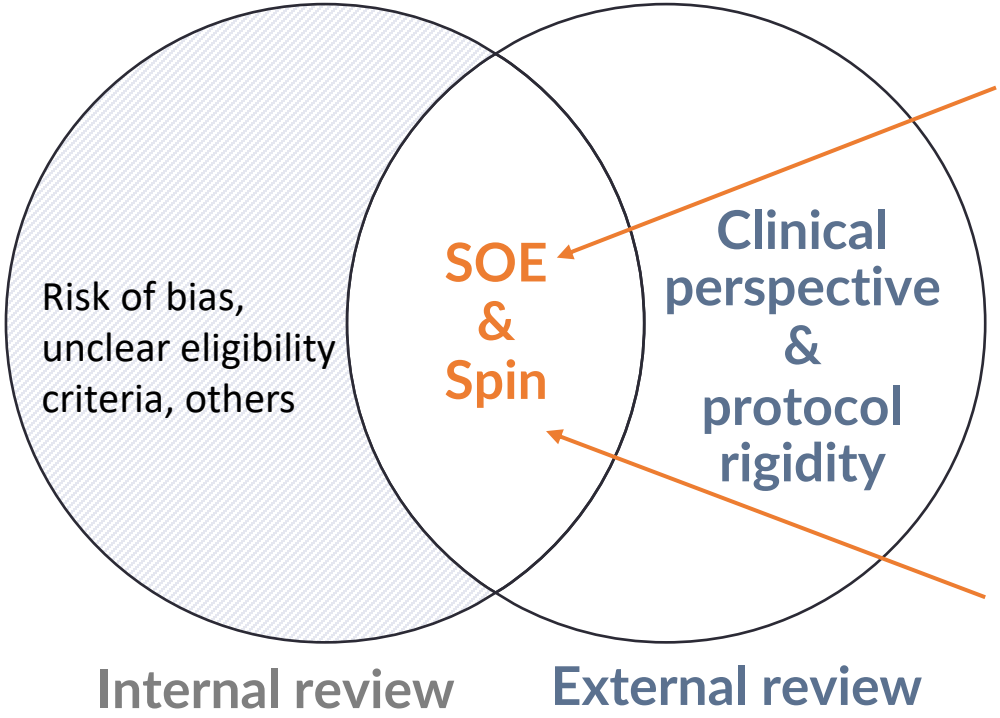


The dominant issue: Nearly half of all substantive feedback concerned the rating of evidence.

External reviewers echoed internal review concerns but added clinical reality.



Concerns about Strength of Evidence and spin persisted



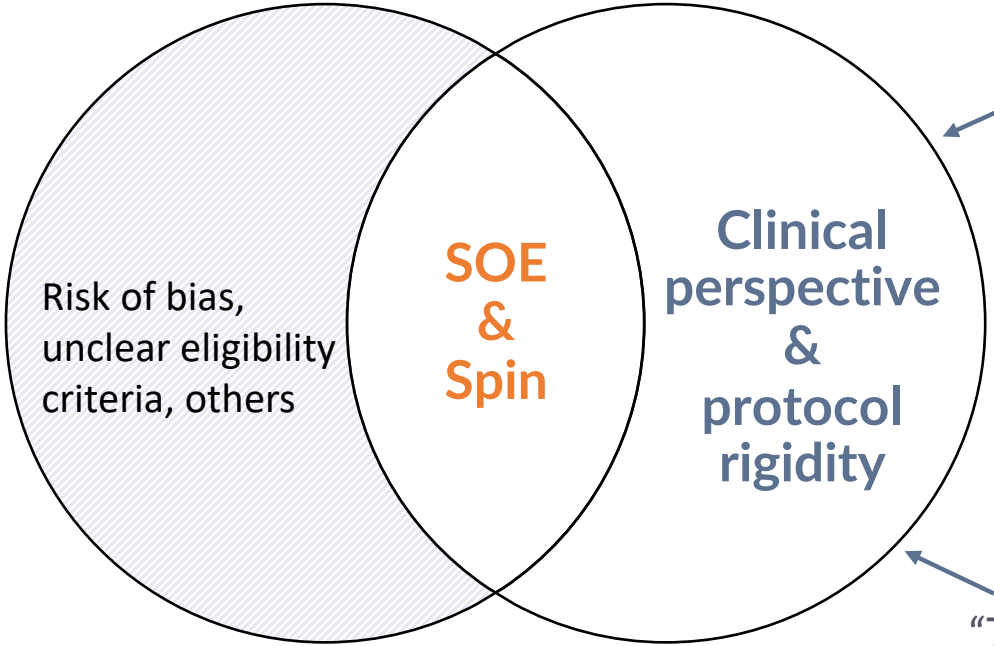
SOE:

- Unclear how ratings were determined
- Inconsistencies (e.g., identical entries for domains yet different ratings; 'moderate' for one outcome and 'low' for another)
- *"Low SOE is very generous."*
- Uncertainty: whether to use qualifying language to indicate SOE or include SOE rating, or both

Spin or subjective interpretation of findings:

- Conclusions too strong, not supported by evidence (e.g., extrapolation)
- Quality of evidence not incorporated into discussion/conclusions
- Biased tone in discussion
- *"Conclusions need tempering."*

What external reviewers added



- Missing clinical expert input
- Certain interventions should have been included; inappropriate comparison group
- Outcomes for SOE grading were determined *a priori*, yet additional ratings would be useful for end-users, guideline developers.

"Strict inclusion criteria is problematic for emerging fields; why not expand criteria to be able to say something?"

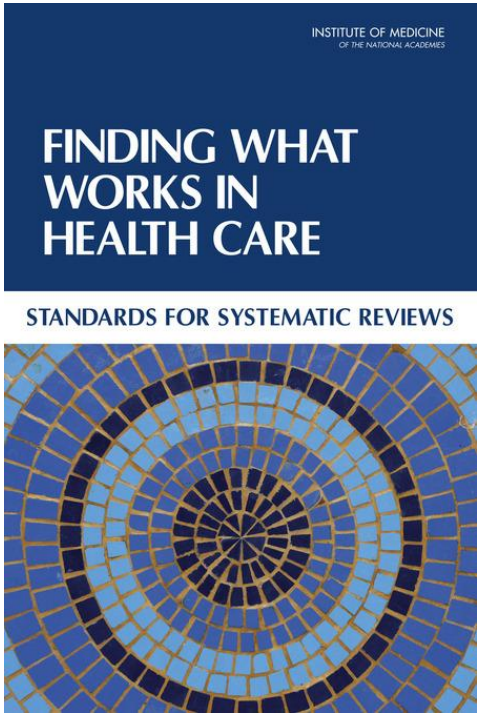
"The rigid approach… results in a report that is likely to be of limited value to the practicing neurological provider."

Conclusions

- We need to revisit how we grade and communicate strength of evidence.
- External reviewers often find clinical value of a report to be limited.
- Strict adherence to the protocol can result in exclusion of relevant/helpful evidence.

Acknowledgements

- Mark Helfand, MD, MPH, MS
- Edi Kuhn, BS, BA
- Celia Fiordalisi, MS
- Camber Hansen-Karr, BA
- Ed Reid, MS, MAT, MFA
- Devan Kansagara, MD, MCR
- Dave Hickam, MD



Implications for the NASEM standards