

Appraisal

Clinimetrics: Appraisal of Guidelines, Research and Evaluation II

Summary

Description: The Appraisal of Guidelines, Research and Evaluation (AGREE) instrument was initially developed in 2003 as a generic tool with which to assess the methodological quality and reporting of clinical practice guidelines.¹ In order to improve the measurement properties of the original instrument, the updated AGREE II tool was more recently developed in 2010, with supporting materials including an updated user manual designed to improve user confidence in application of the AGREE II tool.^{2–4} The AGREE II is extensively used to inform appraisal of guideline quality, policy-related decisions, and to direct guideline development and reporting.^{5–7}

The AGREE II consists of 23 items, which are divided into six domains:²

1. Scope and purpose
2. Stakeholder involvement
3. Rigour of development
4. Clarity of presentation
5. Applicability
6. Editorial independence

Items are rated on a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree), based on criteria and considerations specific to each item. A higher score indicates superior quality.³ Independent assessment is completed by two to four appraisers.² A quality score is calculated for each domain by scaling the sum of all appraisers' relevant item scores as a percentage of the maximum possible score for that domain.⁸ Domain scores can be compared between guidelines and utilised to inform decision-making regarding uptake of the guideline

recommendations.⁸ Users also complete two overall assessments: first, overall quality is rated using the 7-point Likert scale; second, guidelines are either assessed as 'recommended for use' (this may be conditional) or 'not appropriate for use'. No specific quality cut-offs are provided.

The instrument together with online training modules and a detailed user manual are available at no cost from www.agreetrust.org. No specific training or expertise beyond reading the user manual is required for application of the AGREE II, with appraisals taking an average of 90 minutes to complete.⁹ Additionally, the online platform 'My AGREE Plus' supports multiple appraisers to complete blinded quality assessments, and automatically calculates domain scores.

Clinimetric properties: The AGREE II is a reliable tool to be used by a variety of professionals across a range of topic areas. The number of appraisers required to achieve an inter-rater reliability level of 0.7 ranges from two to five across the domains.³ Intraclass correlation averages range from 0.46 to 0.73 for two appraisers, and 0.63 to 0.84 for four appraisers.³ The internal consistency for the domains is moderate to good, with Cronbach's alphas of 0.89, 0.73, 0.75, 0.68, 0.80 and 0.64, respectively.³ Five of the six domains are statistically significant positive predictors of the overall quality assessment rating ($p < 0.001$ to $p < 0.008$).³

Face validity is supported by usefulness scores for each item of 4.80 to 6.53 out of 7.³ Construct validity is demonstrated with 18 out of 21 domains rated significantly higher for high-quality content compared to low-quality content.⁴ This supports the instrument's ability to detect differences in guideline quality.

Commentary

Clinical practice guidelines provide an important avenue for supporting implementation of evidence-based practice. However, the usefulness of clinical practice guidelines is impacted by their quality. With an increasing emergence of published clinical practice guidelines, it is important for clinicians, researchers and policymakers to be able to make sound decisions about which recommendations to implement. Evaluating and comparing clinical practice guideline quality is an important step in this decision-making process, which is supported by the AGREE II instrument.

The utility of the AGREE II tool is supported by its cost-effectiveness, breadth and ease of application. However, the time taken to complete appraisals and the number of concurrent appraisals required to achieve a reliable assessment may impact feasibility. Whilst the AGREE II provides a valid and reliable assessment of the methodological quality of a clinical practice guideline, it does not assess the appropriateness of the clinical content of specific recommendations or the quality of the evidence underlying those recommendations. Further research by the AGREE Enterprise to develop a complementary tool - the AGREE-REX - to evaluate the clinical credibility and implementation of clinical practice guidelines recommendations is currently underway.

Provenance: Invited. Not peer reviewed.

Jenna Lang^{a,b} and Kimberley Haines^{a,b,c}

^aAustralian and New Zealand Intensive Care Research Centre, Monash University

^bDepartment of Physiotherapy, Western Health
^cDepartment of Physiotherapy, The University of Melbourne, Australia

References

1. The AGREE Collaboration. *Qual Saf Health Care*. 2003;12:18–23.
2. Brouwers MC, et al. *CMAJ*. 2010;182:E839–E842.
3. Brouwers MC, et al. *CMAJ*. 2010;182:1045–1052.
4. Brouwers MC, et al. *CMAJ*. 2010;182:E472–E478.
5. Lin I, et al. *Br J Sports Med*. 2018;52:337–343.
6. Chiappini E, et al. *BMJ Open*. 2017;7:e015404.
7. NHMRC; 2018. <https://www.nhmrc.gov.au/guidelinesforguidelines/plan/adopt-adapt-or-start-scratch>. Accessed 25 March, 2019.
8. AGREE Next Steps Consortium. 2013. The AGREE II Instrument [Electronic version]. <https://www.agreetrust.org>. Accessed 25 March, 2019.
9. Brouwers MC, et al. *Implement Sci*. 2011;6:1–10.