

O72

Poor agreement on assessing the quality of clinical practice guidelines in France

Emmanuel Nouyrigat, Joëlle André-Vert, Catherine Solomon-Alexander, Brice Kitio, Patrice Dosquet
National Authority for Health, Saint-Denis, France

Background: Since September 2008 in France, an independent committee of healthcare professionals and patient representatives ("*Comité de validation des recommandations*" (CVR)) critically assesses the final draft of clinical practice guidelines (CPGs) before publication.

Purpose: To estimate the inter-rater reliability among committee members in their assessment of CPG quality.

Methods: All CPGs submitted to the CVR between September 2008 and April 2009 were independently reviewed by at least 2 CVR members using a version of the AGREE instrument adapted to the French context. The instrument consists of 26 items ranked on a 4-point Likert scale. Of these, 25 are organized into 6 domains and one item gives the raters' global judgment on CPG quality. The decision on whether to approve the CPG is taken during a once-monthly meeting after a debate and a vote. Inter-rater reliability was measured by the intra-class correlation (ICC) coefficient and agreement between the raters' global judgment on CPG quality and the final decision by the Kappa coefficient.

Results: A total of 23 CPGs were assessed by the CVR and 12 were approved. The median number of raters per CPG was 4 (range: 2-6). Median inter-rater reliability was poor (ICC= 0.20; interquartile range [IQR]: 0.11 to 0.39) but varied across the domains. Agreement between median overall CPG quality and the final decision was moderate overall (kappa=0.57, 95%CI [0.16 - 0.91]) but good in cases of consensus among raters (n = 17 CPG) (kappa=0.76, 95%CI [0.37 - 1]). When they were no consensus between raters (n=6 CPG), two were approved by the committee and four were rejected.

Discussion: Although the AGREE instrument we used is of value in identifying good quality CPGs, CVR members differed, sometimes substantially, in their assessments. Efforts to reduce variability may help better identify good quality CPGs.