Multiple screening tools, multiple thresholds, multiple clinical cohorts: Evaluating screening tools for obstructive sleep apnoea

Suzanne C Freeman[1], Emer M Brady[2], Helena Polmann[3], Jessica Réus[3], Graziela De Luca Canto[3], Noelle Robertson[4], Iain Squire[5], Lizelle Bernhardt[5,6]

[1] Department of Health Sciences, University of Leicester, Leicester, UK
[2] Diabetes Centre, University Hospitals of Leicester NHS Trust, Leicester, UK
[3] Brazilian Centre for Evidence-based Research/Centro Brasileiro de Pesquisas Baseadas em Evidências, Federal University of Santa Catarina, Florianopólis, Brazil
[4] Department of Neuroscience, Psychology and Behaviour, University of Leicester, Leicester, UK
[5] Department of Cardiovascular Sciences, University of Leicester, Leicester, UK
[6] Community Health Services, Leicestershire Partnership NHS Trust, Leicester, UK

Background: In the UK, 1.4 million people live with undiagnosed and untreated obstructive sleep apnoea (OSA) and are at an increased risk of cardio-metabolic complications and diabetes. Polysomnography (PSG) is the gold standard for the diagnosis of OSA but is expensive, time-consuming and has long waiting lists. A questionnaire to identify patients at high risk of OSA requiring further investigation and treatment would be of great benefit.

Aims: To determine the best questionnaire for identifying adults at high risk of OSA amongst different clinical cohorts accounting for multiple questionnaires and multiple thresholds.

Methods: 31 studies reporting the diagnostic accuracy of the Berlin, STOP or STOP-Bang questionnaires as a screening tool for moderate-to-severe OSA were available for meta-analysis from two clinical cohorts of patients: sleep clinic and surgical. Within each cohort random effects bivariate binomial models were fitted to each questionnaire. Where there was a difference in diagnostic ability between questionnaires we tested this using meta-regression. In the surgical cohort, we accounted for multiple thresholds using the methods of Steinhauser et al[1].

Results: In both the sleep clinic and surgical cohorts, meta-regression including questionnaire as a covariate identified statistical differences in sensitivity between STOP-Bang and Berlin. There was no evidence of differences in specificity. Due to the large number of parameters estimated when accounting for multiple thresholds we were only able to fit two of the eight models proposed by Steinhauser et al[1].

Conclusions: Performing a coherent analysis under the frequentist framework that is able to incorporate multiple questionnaires and multiple thresholds across different clinical cohorts whilst avoiding the well-known issues associated with multiple testing can be challenging within the limits of current methodology, even with a moderately sized dataset.

Keywords
Diagnosis, meta-analysis, screening

References

[1] Steinhauser et al. Modelling multiple thresholds in meta-analysis of diagnostic test accuracy studies. BMC Med Res Meth 2016;16:97