Response to Achieving Consensus in the Measurement of Psychological Adjustment to Cleft Lip and/or Palate at Age 8+ Years

Conrad J. Harrison[1], Louise Dalton[2], Timothy E. E. Goodacre[3,4], and Marc C. Swan[4]

Abstract

In this letter, we discuss the recently published paper by Stock et al, entitled Achieving Consensus in the Measurement of Psychological Adjustment to Cleft Lip and/or Palate at Age 8+ Years.

Keywords

quality of life, psychological assessment, psychosocial adjustment

Dear Editor,

We read with interest this issue's article by (Stock et al., 2020), which reviewed patient-reported measures for psychological adjustment to cleft lip and/or palate in people aged 8 years or older. The authors aimed to appraise available measures, before recommending a core set, presumably for use in both research and clinical practice. We commend the authors for their focus on this issue, which is as important as it is challenging. However, we have some reservations about this article's methodology and reporting.

In 2017, this journal published an article from an International Consortium for Health Outcomes Measurement (ICHOM) working group that aimed to achieve consensus on measurement across many domains of cleft-related health (Allori et al., 2017). Stock et al may have good reason to repeat part of this process, but this was not explained in their article. A critical discussion of the limitations of ICHOM's core outcome set would have been interesting and could have helped to contextualize this work.

The authors included measures that had at least "acceptable" levels of validity and reliability. However, what constitutes "acceptable" was not stated. While the authors did report whether studies of a measure's validity existed, they made no attempt to quantify, qualify, or compare the level of evidence that supports each measure. The authors include the About You (CEN-Q) measure in their recommended set, although according to their paper, there is no evidence to support its validity or reliability. We would like to remind readers that validity is not a binary concept. It is more helpful to qualify the evidence that

implies an instrument's validity than to describe a measure as valid or invalid.

Similarly, normative data are population-specific and details about the normative samples referred to in this paper would help readers to understand why the authors considered them translatable. Ideally, a core measurement set would be of use internationally.

Assessing the quality of measurement instruments is a complex process. However, the consensus-based standards for the selection of health status measurement instruments (COSMIN) does provide comprehensive and robust guidance on conducting systematic reviews of measurement instruments (Mokkink et al., 2018). We would be interested to know why the authors chose not to follow this widely accepted guideline for the methodology and reporting of such studies.

Subscales of the CLEFT-Q were deemed too long to include in the author's core set, and we found this decision confusing. The CLEFT-Q is a modular measure. It is made up of 13 questionnaires which range from 6 to 12 items in length

[1] Nuffield Department of Orthopedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK
[2] Department of Psychiatry, University of Oxford, Oxford, UK
[3] Royal College of Surgeons of England, London, UK
[4] Spires Cleft Centre, John Radcliffe Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

Corresponding Author:
Conrad J. Harrison, The Botnar Research Centre, University of Oxford, Old Road, Headington, Oxford, OX3 7LD, UK.
Email: conrad.harrison@medsci.ox.ac.uk

(Klassen et al., 2018). Each questionnaire functions independently, so the assessor can choose which questionnaires are relevant to the person they are assessing. The development of the CLEFT-Q computerized adaptive test is likely to reduce this response burden even further (Harrison et al., 2019).

Moving forward, we think that consensus in this subject can only be achieved through collaboration. We are concerned that researchers in this field are working independently and in parallel, rather than synergistically. Stock et al have gone to considerable effort to involve patients in their work, and this is highly admirable. But consensus will not be achieved unless a critical mass of stakeholders are involved, and this includes other health care providers and researchers from the international cleft community.

Declaration of Conflicting Interests

Funding

References

Allori AC, Kelley T, Meara JG, Albert A, Bonanthaya K, Chapman K, Cunningham M, Daskalogiannakis J, de Gier H, Heggie AA, et al. A standard set of outcome measures for the comprehensive appraisal of cleft care. In: *Cleft Palate-Craniofacial J.* 2017;54(5): 540-554. doi:10.1597/15-292

Harrison CJ, Geerards D, Ottenhof MJ, Klassen AF, Wong Riff KWY, Swan MC, Pusic AL, Sidey-Gibbons CJ. Computerised adaptive testing accurately predicts CLEFT-Q scores by selecting fewer, more patient-focused questions. *J Plast Reconstr Aesthetic Surg.* 2019;72(11):1819-1824. doi:10.1016/j.bjps.2019.05.039

Klassen AF, Wong Riff KWY, Longmire NM, Albert A, Allen GC, Asim Aydin M, Baker SB, Cano SJ, Chan AJ, Courtemanche DJ, et al. Psychometric findings and normative values for the CLEFT-Q based on 2434 children and young adult patients with cleft lip and/or palate from 12 countries. *CMAJ.* 2018;190(5): E455-E462. doi:10.1503/cmaj.170289

Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, Terwee CB. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res.* 2018;27(5):1171-1179. doi:10.1007/s11136-017-1765-4

Stock NM, Hammond V, Hearst D, Owen T, Edwards Z, Ridley M, Rumsey N. Achieving consensus in the measurement of psychological adjustment to cleft lip and/or palate at age 8+ years. *Cleft Palate-Craniofacial J.* 2020:57(6):746-752. doi:10.1177/ 1055665619898596