Quality of Orthopaedic Patient-Reported Outcome Measure Validation Studies is Often Insufficient

Ville T Ponkilainen, Ilari Kuitunen, Aarni Oskari Pakarinen, Matias Aleksanteri Vaajala, Rasmus Juhani Liukkonen, Hannes Taavetti Tukiainen, Mikko Uimonen

INTRODUCTION:
Validation studies, i.e., studies examining validity of psychometric measurement properties of PROMs, has developed into whole new field in clinical research. To ensure the high quality in validation research, recommendations, and guidelines have been developed by the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) steering committee. The COSMIN checklist includes point-by-point standards for adequate quality assessment of the main dimensions of PROM validity. Nevertheless, the quality of validation studies has frequently been questioned.
The aim of this study was to examine the conclusions of the patient-reposted outcome measure (PROM) validation studies and the methodological basis behind these conclusions.
METHODS:
This systematic review was performed on studies evaluating psychometric properties of PROMs used in orthopaedic surgery. Studies published between 1 June and 31 December 2021 were identified from Web of Science (Clarivate) and Scopus (Elsevier) databases. The quality of validity subfield evaluation in the studies was assessed according to COSMIN checklist. Nine of the validity subfields were assessed.
RESULTS:
In the 59 included studies, median sample size was 111 (IQR 97 – 204) and 19 (32%) of the studies had insufficient sample size according to the COSMIN checklist. Of nine validity subfields, mean number of properly assessed subfields was 3.7 (SD 1.6). In 50 (85%) of the studies, the conclusion was phrased deterministically that the PROM is “valid.” In these studies, the mean number of evaluated validity subfields was 3.9 (SD 1.5) out of nine. None of the studies reported that the PROM is “not valid.”
DISCUSSION AND CONCLUSION:
The methodological basis of the conclusions drawn in the studies investigating psychometric properties of orthopaedic PROMs are often insufficient. The studies are often performed with too small sample sizes and focus on only few validity subfields and still present deterministic conclusions that a PROM is “valid.”