P92 Associated patient-reported outcome measures: Do linked change scores tell the whole story?

Daniel Deutscher[1,2,3], Maxwell Armand Mansolf[4], Benjamin D Schalet[5], David Cella[4]

[1]Net Health Systems, Pittsburgh, USA. [2]Maccabi Healthcare Services, Tel Aviv, Israel. [3]University of Utah, Salt Lake City, USA. [4]Northwestern University Feinberg School of Medicine, Chicago, USA. [5]Department of Epidemiology and Data Science, Amsterdam UMC, Amsterdam, Netherlands

Objective: Scores from patient-reported outcome measures (PROMs) that are conceptually similar and sufficiently correlated can be linked using a reliable linking function. Linked scores can also be used longitudinally to assess change. We aimed to assess whether commonly used PROMs of functional status produce biased linked change scores and explore possible solutions if such bias exists.

Methods: Patients treated for low back pain completed the FOTO Lumbar Computerized Adaptive Test (LCAT) and the PROMIS® Physical Function CAT v2.0 (PFCAT) at the start of physical therapy. Equipercentile score linking was applied. A subset of patients also completed the same PROMs at discharge, allowing computation of LCAT (primary) and linked (PFCAT to LCAT) change scores. Two sets of primary and linked predicted change scores were calculated by applying a risk-adjustment model using a set of independent covariates including either the LCAT or the linked baseline scores. This resulted in two sets of primary and linked residual scores, defined as the difference between observed and predicted change scores. Unadjusted and adjusted linked residuals were compared to assess bias.

Results: Patients from a development sample (n=3446, mean age=57) had baseline LCAT and PFCAT mean (SD; range) scores of 51.0 (13.6; 1 to 94) and 39.5 (7.2; 19 to 76), respectively. The two sets of scores were highly correlated (r=0.78), and had similar distributions, confirming suitability for score linking. Consistent results with a validation sample (n=862) supported the external validity of the linking function. For patients with scores at discharge (n=2,363), observed change was 12.6 for the LCAT, and 8.9 for the LCAT scores linked from PFCAT scores. The lower mean residual score for the linked PFCAT scores (-4.0) compared to the LCAT mean residual (-0.2), suggest that linking resulted in biased estimates of linked residual scores. A second equipercentile linking function of the residuals eliminated most of the bias.

Conclusions: Results demonstrated biased linked change scores, despite conceptually similar item content. To adjust for differences in change scores, we applied a second linking function to residuals. Other correction methods such as statistical adjustment based on the reliability of change scores should also be considered.