

O117 Novel approach to reduce item burden in longitudinal PROMIS-CAT Assessment: Simulation using PROMIS Pain Interference

Jad Fadlallah, Devron Swaby, Hayden John, Mowa Ayibiowu, Ward Hajjar, Nathaniel Edwards, Nishat Syeda, Mousa El-Sururi, Istvan Mucsi

Ajmera Transplant Centre and Division of Nephrology, University Health Network, Toronto, Canada

Objective: Computer Adaptive Testing (CAT) offers precise assessment of patient-reported outcomes. However, the perceived burden of repeated questionnaire completion in longitudinal settings remains a challenge. We propose a novel *Longitudinal CAT* (LT-CAT) algorithm to reduce item burden, which we evaluated through a simulation using existing PROMIS Pain Interference CAT (PI-CAT) T-scores.

Methods: Longitudinal convenience sample of adult kidney, kidney-pancreas, and liver transplant recipients (≥ 60 days post-transplant), who completed PROMIS PI-CAT (higher score=worse symptoms) via an electronic data capture system at baseline (week-0) and follow-up (week-12).

In CAT, the final item is selected for providing maximum information at the current estimated latent trait. We simulated a scenario, where the last item answered at baseline would have been the first item administered at follow-up. If the response remained unchanged, the baseline score was carried forward to week-12 =LT-CAT score. To assess accuracy of LT-CAT compared to CAT at follow-up, we calculated the mean absolute error(MAE), root mean square error(RMSE), and agreement in symptom severity categories (<60 =none/mild, ≥ 60 =moderate/severe). To evaluate group-level agreement, we also calculated Cohen's d and Pearson's r (95% CI). We also simulated screen-to-CAT(ST-CAT) scores and report the total number of items that would have been administered using LT-CAT, and ST-CAT compared to CAT.

Results: Of 115 participants completing both baseline and follow-up, LT-CAT scores were simulated for 59(51%) of participants. For this subset of participants, MAE was 1.81(95% CI: 1.03;2.59), RMSE was 3.57(95% CI: 2.46;4.68) Only 2(3%) patients were misclassified with LT-CAT (CAT \rightarrow LT-CAT scores were: 57.72 \rightarrow 60.54; 60.43 \rightarrow 58.46). Effect size was negligible (Cohen's d: -0.10 [95% CI: -0.36 ;0.15]) and correlation was strong (Pearson's r=0.92[95% CI: 0.86;0.95]).

At week-12, 115 participants completed 819 items with CAT. They would have completed 361 with LT-CAT, 295 with ST-CAT and 232 with the combined use of both LT-CAT and ST-CAT.

Conclusion: LT-CAT offers a strategy for reducing respondent burden in longitudinal studies with minimal loss in accuracy. In fact, LT-CAT and ST-CAT can be combined, to further reduce respondent burden.