

Schalet, B. D., Lim, S., Cella, D., Choi, S. W. Linking scores with patient-reported health outcome instruments: A validation study and comparison of three linking methods. *Psychometrika*.

Online Resource 1: Summary of Linking Studies with PROMIS and other Patient-Reported Outcome Instruments

Published reports on patient-reported outcome (PRO) instrument linking have been typically been organized around a single health domain. Linking studies with PROMIS instruments (Cella et al., 2010) cover the domains of depression (Choi, Schalet, Cook, & Cella, 2014; Gibbons et al., 2011; Kim et al., 2015; Olino et al., 2013; Schalet et al., 2016; Schalet et al., 2020; Wahl et al., 2014), pediatric depression (Kaat et al., 2020), anxiety (Schalet, Cook, Choi, & Cella, 2014; Victorson et al., 2019), fatigue (Lai, Cella, Yanez, & Stone, 2014; Noonan et al., 2012), physical function (Kaat, Schalet, Rutsohn, Jensen, & Cella, 2018; Schalet, Revicki, et al., 2015; Voshaar et al., 2019), pain interference (Askew et al., 2013; Chen, Revicki, Lai, Cook, & Amtmann, 2009; Cook, Schalet, Kallen, Rutsohn, & Cella, 2015), and general or global health (Schalet, Rothrock, et al., 2015). A few of these studies linked a large number of instruments, of which PROMIS is one, to a larger metric (Voshaar et al., 2019; Wahl et al., 2014). PROMIS linking studies have also been used to facilitate score translation from pediatric to adult measures for emotional distress (Reeve et al., 2016) and physical health domains (Tulsky et al., 2019). Using regression techniques, some researchers have linked PROMIS-based multidimensional health status scores, which are useful for economic evaluations (Feeny, Furlong, Boyle, & Torrance, 1995). These scores been linked using regression techniques (Hays et al., 2016; Revicki et al., 2009; Thompson, Lapin, & Katzan, 2017), noting that care needs to be taken to avoid regression to the mean (Fayers & Hays, 2014).

Some linking studies incorporated PROMIS instruments, but did not use the established parameters (available via help@healthmeasures.net for research purposes). PROMIS item parameters were not used in co-calibration studies with on a large number of PROs, of which PROMIS is only one (Voshaar et al., 2019; Wahl et al., 2014), or when the research question does not require anchoring (Olino et al., 2013). Linking analyses on a new (freely estimated) metric may have the advantage of producing more accurate standard errors and are less likely to result in bias; however, such an approach carries the practical risk for PROMIS of introducing and managing multiple calibration sets. This may be minimized by publicly cataloging the calibrations and facilitating scoring (Fischer & Rose, 2016).

PRO-based linking studies *without* PROMIS instruments have included instruments that measure social anxiety (Sunderland et al., 2018), psychological distress (Batterham, Sunderland, Slade, Callear, & Carragher, 2018), generalized anxiety (Kisala et al., 2015), depression (Fischer, Tritt, Klapp, & Fliege, 2011; Orlando, Sherbourne, & Thissen, 2000; Tulsy et al., 2015), disruptive behavior (Kaat et al., 2019), personality disorder severity (Zimmermann et al., 2020), fatigue (Friedrich et al., 2019), physical function or functional status (McHorney & Cohen, 2000; Shoop-Worrall et al., 2020; ten Klooster et al., 2013; Voshaar et al., 2019), as well as multi-dimensional scales (Holzner et al., 2006).

References

- Askew, R. L., Kim, J., Chung, H., Cook, K. F., Johnson, K. L., & Amtmann, D. (2013). Development of a crosswalk for pain interference measured by the BPI and PROMIS pain interference short form. *Quality of Life Research*, 22(10), 2769-2776.
- Batterham, P., Sunderland, M., Slade, T., Callear, A., & Carragher, N. (2018). Assessing distress in the community: psychometric properties and crosswalk comparison of eight measures of psychological distress. *Psychological medicine*, 48(8), 1316-1324.
- Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., . . . Choi, S. (2010). The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and

- tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of clinical epidemiology*, 63(11), 1179-1194.
- Chen, W.-H., Revicki, D. A., Lai, J.-S., Cook, K. F., & Amtmann, D. (2009). Linking pain items from two studies onto a common scale using item response theory. *Journal of pain and symptom management*, 38(4), 615-628.
- Choi, S. W., Schalet, B., Cook, K. F., & Cella, D. (2014). Establishing a common metric for depressive symptoms: Linking the BDI-II, CES-D, and PHQ-9 to PROMIS Depression. *Psychological assessment*, 26(2), 513.
- Cook, K. F., Schalet, B. D., Kallen, M. A., Rutsohn, J. P., & Cella, D. (2015). Establishing a common metric for self-reported pain: Linking BPI pain interference and SF-36 bodily pain subscale scores to the PROMIS pain interference metric. *Quality of Life Research*, 24(10), 2305-2318.
- Fayers, P. M., & Hays, R. D. (2014). Should linking replace regression when mapping from profile-based measures to preference-based measures? *Value in Health*, 17(2), 261-265.
- Feeny, D., Furlong, W., Boyle, M., & Torrance, G. W. (1995). Multi-attribute health status classification systems. *Pharmacoeconomics*, 7(6), 490-502.
- Fischer, F., & Rose, M. (2016). www.common-metrics.org: a web application to estimate scores from different patient-reported outcome measures on a common scale. *BMC medical research methodology*, 16(1), 142.
- Fischer, H. F., Tritt, K., Klapp, B. F., & Fliege, H. (2011). How to compare scores from different depression scales: Equating the Patient Health Questionnaire (PHQ) and the ICD-10-Symptom Rating (ISR) using item response theory. *International journal of methods in psychiatric research*, 20(4), 203-214.
- Friedrich, M., Hinz, A., Kuhnt, S., Schulte, T., Rose, M., & Fischer, F. (2019). Measuring fatigue in cancer patients: a common metric for six fatigue instruments. *Quality of Life Research*, 28(6), 1615-1626.
- Gibbons, L. E., Feldman, B. J., Crane, H. M., Mugavero, M., Willig, J. H., Patrick, D., . . . Crane, P. K. (2011). Migrating from a legacy fixed-format measure to CAT administration: calibrating the PHQ-9 to the PROMIS depression measures. *Quality of Life Research*, 20(9), 1349-1357.
- Hays, R. D., Revicki, D. A., Feeny, D., Fayers, P., Spritzer, K. L., & Cella, D. (2016). Using linear equating to map PROMIS® global health items and the PROMIS-29 V2. 0 profile measure to the health utilities index mark 3. *Pharmacoeconomics*, 34(10), 1015-1022.
- Holzner, B., Bode, R., Hahn, E. A., Cella, D., Kopp, M., Sperner-Unterweger, B., & Kemmler, G. (2006). Equating EORTC QLQ-C30 and FACT-G scores and its use in oncological research. *European Journal of Cancer*, 42(18), 3169-3177.
- Kaat, A. J., Blackwell, C. K., Estabrook, R., Burns, J. L., Petitclerc, A., Briggs-Gowan, M. J., . . . Wakschlag, L. S. (2019). Linking the Child Behavior Checklist (CBCL) with the Multidimensional Assessment Profile of Disruptive Behavior (MAP-DB): Advancing a dimensional spectrum approach to disruptive behavior. *Journal of child and family studies*, 28(2), 343-353.
- Kaat, A. J., Kallen, M. A., Nowinski, C. J., Sterling, S. A., Westbrook, S. R., & Peters, J. T. (2020). PROMIS® pediatric depressive symptoms as a harmonized score metric. *Journal of Pediatric Psychology*, 45(3), 271-280.
- Kaat, A. J., Schalet, B. D., Rutsohn, J., Jensen, R. E., & Cella, D. (2018). Physical function metric over measure: An illustration with the Patient-Reported Outcomes Measurement

- Information System (PROMIS) and the Functional Assessment of Cancer Therapy (FACT). *Cancer*, 124(1), 153-160.
- Kim, J., Chung, H., Askew, R. L., Park, R., Jones, S. M., Cook, K. F., & Amtmann, D. (2015). Translating CESD-20 and PHQ-9 Scores to PROMIS Depression. *Assessment*, 1073191115607042.
- Kisala, P. A., Tulsy, D. S., Kalpakjian, C. Z., Heinemann, A. W., Pohlig, R. T., Carle, A., & Choi, S. W. (2015). Measuring anxiety after spinal cord injury: Development and psychometric characteristics of the SCI-QOL Anxiety item bank and linkage with GAD-7. *The journal of spinal cord medicine*, 38(3), 315-325.
- Lai, J.-S., Cella, D., Yanez, B., & Stone, A. (2014). Linking fatigue measures on a common reporting metric. *Journal of pain and symptom management*, 48(4), 639-648.
- McHorney, C. A., & Cohen, A. S. (2000). Equating health status measures with item response theory: illustrations with functional status items. *Medical care*, II43-II59.
- Noonan, V. K., Cook, K. F., Bamer, A. M., Choi, S. W., Kim, J., & Amtmann, D. (2012). Measuring fatigue in persons with multiple sclerosis: creating a crosswalk between the Modified Fatigue Impact Scale and the PROMIS Fatigue Short Form. *Quality of Life Research*, 21(7), 1123-1133.
- Olino, T. M., Yu, L., McMakin, D. L., Forbes, E. E., Seeley, J. R., Lewinsohn, P. M., & Pilkonis, P. A. (2013). Comparisons across depression assessment instruments in adolescence and young adulthood: an item response theory study using two linking methods. *Journal of abnormal child psychology*, 41(8), 1267-1277.
- Orlando, M., Sherbourne, C. D., & Thissen, D. (2000). Summed-score linking using item response theory: application to depression measurement. *Psychological assessment*, 12(3), 354.
- Reeve, B. B., Thissen, D., DeWalt, D. A., Huang, I.-C., Liu, Y., Magnus, B., . . . Ni, P. (2016). Linkage between the PROMIS® pediatric and adult emotional distress measures. *Quality of Life Research*, 25(4), 823-833.
- Revicki, D. A., Kawata, A. K., Harnam, N., Chen, W.-H., Hays, R. D., & Cella, D. (2009). Predicting EuroQol (EQ-5D) scores from the patient-reported outcomes measurement information system (PROMIS) global items and domain item banks in a United States sample. *Quality of Life Research*, 18(6), 783-791.
- Schalet, B., Kaat, A., Buckenmaier III, C., Barnhill, R., Vrahas, M., & Gershon, R. (2016). *Extending the ceiling of an item bank: Development of above-average physical function items for PROMIS*. Paper presented at the International Society for Quality of Life Research, Copenhagen, Denmark.
- Schalet, B. D., Cook, K. F., Choi, S. W., & Cella, D. (2014). Establishing a common metric for self-reported anxiety: linking the MASQ, PANAS, and GAD-7 to PROMIS Anxiety. *Journal of anxiety disorders*, 28(1), 88-96.
- Schalet, B. D., Janulis, P., Kipke, M. D., Mustanski, B., Shoptaw, S., Moore, R., . . . Ragsdale, A. (2020). Psychometric Data Linking Across HIV and Substance Use Cohorts. *AIDS AND BEHAVIOR*.
- Schalet, B. D., Revicki, D. A., Cook, K. F., Krishnan, E., Fries, J. F., & Cella, D. (2015). Establishing a common metric for physical function: Linking the HAQ-DI and SF-36 PF subscale to PROMIS® Physical Function. *Journal of general internal medicine*, 30(10), 1517-1523.

- Schalet, B. D., Rothrock, N. E., Hays, R. D., Kazis, L. E., Cook, K. F., Rutsohn, J. P., & Cella, D. (2015). Linking Physical and Mental Health Summary Scores from the Veterans RAND 12-Item Health Survey (VR-12) to the PROMIS® Global Health Scale. *Journal of general internal medicine, 30*(10), 1524-1530.
- Shoop-Worrall, S. J., Oude Voshaar, M. A., McDonagh, J. E., Van de Laar, M. A., Wulffraat, N., Thomson, W., . . . Verstappen, S. M. (2020). A Common Functional Ability Score for Young People with Juvenile Idiopathic Arthritis. *Arthritis Care & Research*.
- Sunderland, M., Batterham, P., Cleave, A., Carragher, N., Baillie, A., & Slade, T. (2018). High agreement was obtained across scores from multiple equated scales for social anxiety disorder using item response theory. *Journal of clinical epidemiology, 99*, 132-143.
- ten Klooster, P. M., Voshaar, M. A. O., Gandek, B., Rose, M., Bjorner, J. B., Taal, E., . . . van de Laar, M. A. (2013). Development and evaluation of a crosswalk between the SF-36 physical functioning scale and Health Assessment Questionnaire disability index in rheumatoid arthritis. *Health and quality of life outcomes, 11*(1), 1.
- Thompson, N. R., Lapin, B. R., & Katzan, I. L. (2017). Mapping PROMIS global health items to EuroQol (EQ-5D) utility scores using linear and equipercentile equating. *Pharmacoeconomics, 35*(11), 1167-1176.
- Tulsky, D. S., Kisala, P. A., Boulton, A. J., Jette, A. M., Thissen, D., Ni, P., . . . Mulcahey, M. (2019). Determining a transitional scoring link between PROMIS® pediatric and adult physical health measures. *Quality of Life Research, 28*(5), 1217-1229.
- Tulsky, D. S., Kisala, P. A., Kalpakjian, C. Z., Bombardier, C. H., Pohlig, R. T., Heinemann, A. W., . . . Choi, S. W. (2015). Measuring depression after spinal cord injury: Development and psychometric characteristics of the SCI-QOL Depression item bank and linkage with PHQ-9. *The journal of spinal cord medicine, 38*(3), 335-346.
- Victorson, D., Schalet, B. D., Kundu, S., Helfand, B. T., Novakovic, K., Penedo, F., & Cella, D. (2019). Establishing a common metric for self-reported anxiety in patients with prostate cancer: Linking the Memorial Anxiety Scale for Prostate Cancer with PROMIS Anxiety. *Cancer, 125*(18), 3249-3258.
- Voshaar, M. O., Vonkeman, H., Courvoisier, D., Finckh, A., Gossec, L., Leung, Y., . . . Wulffraat, N. (2019). Towards standardized patient reported physical function outcome reporting: linking ten commonly used questionnaires to a common metric. *Quality of Life Research, 28*(1), 187-197.
- Wahl, I., Löwe, B., Bjorner, J. B., Fischer, F., Langa, G., Voderholzer, U., . . . Rose, M. (2014). Standardization of depression measurement: a common metric was developed for 11 self-report depression measures. *Journal of clinical epidemiology, 67*(1), 73-86.
- Zimmermann, J., Müller, S., Bach, B., Hutsebaut, J., Hummelen, B., & Fischer, F. (2020). A common metric for self-reported severity of personality disorder. *Psychopathology, 53*(3), 161-171.