

## Measuring pediatric patient-reported outcomes: good progress but a long way to go

I-Chan Huang · Dennis A. Revicki ·  
Carolyn E. Schwartz

Published online: 22 December 2013  
© Springer Science+Business Media Dordrecht 2014

The pre-Socratic Greek philosopher Protagoras (490 BC–420 BC) made an argument about 2,500 years ago: *Man is the measure of all things—of things which are, how they are, and of things which are not, how they are not* [1]. This controversial idea is in contrast with a long-standing paradigm claiming that the universe is based on objective matters and outside the human's subjective influence. Similarly, the use of objective methods such as laboratory tests and diagnostic imaging rather than subjective and self-reported methods to evaluate patient health outcomes has dominated medical sciences for several 100 years. Although clinicians and policy makers have gradually accepted the notions “quality of life (QOL),” “health-related quality of life (HRQOL),” or “patient-reported outcomes (PROs)” as an endpoint in the adult setting, the pediatric setting has lagged behind in considering self-reports in concept initiation, instrument development, and clinical application of PROs. We conventionally believed that children's self-reported health information is unreliable and together with complex developmental issues, resulted in insufficient attention to and under-utilization of

pediatric PRO data. Indeed, children are able to report their health status in an adequate manner [2], and pediatric PRO instruments need to carefully accommodate specific content related to children's cognitive development as well as reading ability, vocabulary, and language skill.

This special section of *Quality of Life Research* includes four articles exclusively focusing on the theme of pediatric PROs. Two articles applied the International Classification of Functioning, Disability, and Health—Children and Youth Version (ICF-CY) as a framework to compare the conceptual content of pediatric PROs embedded in existing instruments for children and adolescents with chronic conditions [3, 4]. One article synthesized the constructs of PRO instruments for pediatric cancer [5], and another article summarized the development of the KIDSCREEN which is commonly used in European pediatric studies [6]. We applaud the authors for their efforts and useful findings in pediatric PRO measurement.

The comparison among pediatric PRO instruments is a challenging task because these instruments were not created based on the same conceptual foundation. Instrument developers use different conceptual frameworks to generate target PROs with specific content and domains and employ various stakeholder-engagement processes to create items. They also apply different psychometric methods to evaluate PRO measures. Previous review articles on pediatric PRO instruments might be misleading because the comparison is confined to the “face value” of the instrument domains [7–9]. These studies often assumed that the domains named by individual instruments (e.g., physical functioning) reflected the genuine concepts of PROs, neglecting the actual content of the items embedded in each instrument. Consequently, the same domain name across different instruments does not necessarily mean they

---

I.-C. Huang (✉) · D. A. Revicki  
Department of Health Outcomes and Policy, Institute for Child  
Health Policy, College of Medicine, University of Florida,  
PO Box 100177, Gainesville, FL 32610, USA  
e-mail: ichuang@ufl.edu

D. A. Revicki  
Outcomes Research, Evidera, Bethesda, MD, USA

C. E. Schwartz  
DeltaQuest Foundation, Inc, Concord, MA, USA

C. E. Schwartz  
Departments of Medicine and Orthopaedic Surgery, Tufts  
University Medical School, Boston, MA, USA

measure the same concepts. For example, a recent pediatric study utilizing a sample of approximately 1,000 participants found that although four pediatric generic instruments (the Child Health and Illness Profile, the KIDSCREEN-52, the KINDL, and the Pediatric Quality of Life Inventory) use a similar label for physical and psychological domains, the correlations between physical domains of the four instruments were equally the same when compared with the correlation of the physical and psychological domains [10]. This finding suggests poor convergent/discriminant validity and evokes a serious question: what are the elements captured by our pediatric PRO instruments? In this special section, Ahuja et al. [4] concluded that the majority of pediatric PRO instruments for obesity primarily capture the concepts of functional status rather than HRQOL or QOL. The unique contribution of this study is the application of the ICF-CY as a framework to level the field for a systematic, head-to-head comparison among the item content of different instruments. Interestingly, analyzing a clinical trial registration database, Fayed et al. [3] found that non-drug interventions or late phase clinical trials (phase IV) were more likely to include activity and participation as endpoints than drug interventions or earlier phases (phase II or III) clinical trials.

It is encouraging that the two articles using ICF-CY [3, 4] aim to apply universal comparisons of different PRO instruments. Several methodological issues require our careful attention nonetheless. First, there continues to be confusion about the concepts of functional status, HRQOL, and QOL. Although no gold standard definitions are available for these concepts, Wilson and Cleary have proposed a comprehensive model illuminating that HRQOL is a broad concept which comprises three key components: functional status, general health perception, and QOL. However, the term “HRQOL” used by the above-mentioned two articles actually captures the notion “general health perception” that was described in Wilson and Cleary’s classical model [11]. Second, the two articles merely restrict themselves to mapping items of pediatric PROs to different categories (or codes) of ICF-CY. Although this is a useful approach for demonstrating the content validity of PRO tools, empirical data have yet to be collected to confirm the qualitative comparisons. Modern test theory [e.g., item response theory (IRT)] serves as a foundation to assess measurement properties of instruments at item and domain levels. Evidence of IRT parameters on the items assigned to the same category of ICF-CY (e.g., ICF-CY code b152: emotional functions) helps to inform whether these items truly measure different levels of the same underlying latent construct (i.e., emotional function). Combining ICF-CY with IRT, methodologies will empower the researchers to establish item banks with content-valid and psychometrically robust properties.

The article by Anthony et al. [5] specifically compared the content of pediatric PRO instruments designed for children with cancer. In this study, authors adopted an inductive reasoning approach to summarize the content of domains/subdomains from existing generic and cancer-specific tools. They created a working PRO model with four domains for children with cancer: physical health, psychological health, social health, and general health. The resulting four-domain solution is not surprising because it represents the essential constructs of the existing instruments. The question was, are these domains comprehensive enough to capture PROs important to cancer patients in general and to survivors in particular? Given that the survival rate is greater than 80 %, pediatric cancer patients are facing many life challenging issues. For example, one recent article discusses the potential missing content of PROs critical to childhood cancer survivors by reviewing the published “qualitative” studies and comparing the results to the conventional PRO framework with physical, psychological, and social domains [12]. This study found that childhood survivors are behind in several developmental milestones, and survivors have unique concerns related to cancer such as normalcy, independency, fertility, self-identity, and body image. However, these issues are either neglected or simply captured by the inclusion of only a single item in the existing PRO tools [12]. Additionally, as emphasized by Anthony et al. [5], measuring symptom impact and burden are important concepts in pediatric cancer research, yet symptoms are largely overlooked in existing PRO tools. Symptoms are often graded by clinicians rather than self-reported by patients [e.g., National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE)]. Recent evidence suggests that prevalence of symptoms increases significantly alongside the aging process in childhood cancer survivors, and symptom presence explains about 60 % of variance of physical and mental aspects of HRQOL [13].

Ravens-Sieberer et al. [6] summarized the development and psychometric evaluation over the past 15 years for the KIDSCREEN. The popularity of the KIDSCREEN in Europe is parallel to that of the PedsQL, which is commonly used in the United States and other countries [14]. The KIDSCREEN was developed on the basis of qualitative methods for content identification and item generation, followed by quantitative methods (classical test theory and IRT) to generate short and long forms and a utility index for different research purposes. The uniqueness of KIDSCREEN is the inclusion of the domains related to children’s psycho-social-behavioral development such as autonomy, social acceptance (bullying), and the new initiative of a computerized adaptive test module—the KIDSCAT. A recent national survey of pediatricians and specialists suggests that the most beneficial solution for

increasing PRO application in clinical practice is to use CAT equipped with programming to automatically calculate individual's PRO scores and to provide guidance for interventions and follow-up [15].

In contrast to four decades ago, when the term “quality of life” first appeared in the pediatric literature [16], researchers have taken a major step forward to improving pediatric PRO measurement and application. However, in celebrating this achievement, we must also focus on the next best steps for advancing pediatric PRO measurement in the coming decade. First, there is a great need to explore the issues related to the changes in pediatric PRO scores using longitudinal study designs. These issues include responsiveness and minimally important differences (MIDs). Responsiveness methods focus on the change in PRO scores when there is a true change in individual's health status even if the magnitude of the change of PRO scores is small. In contrast, MIDs quantify the magnitude of the change in PRO scores that is clinically meaningful or interpretable using appropriate anchors and statistical methods [17]. Without information on responsiveness and MIDs, we are not able to monitor and interpret PROs over disease or developmental trajectories. Another important but neglected topic is the potential response shift—children and adolescents might change their internal standard to interpret the meaning of PROs differently along with the experience of significant events (e.g., cancer) or across different neurocognitive development stages. Schwartz et al. [18] found that response shift does exist in childhood cancer survivors and that this phenomenon tends to overestimate treatment effects in a shorter-term (1 week) and underestimate treatment effects in a longer-term (3 month) period. Unfortunately, the prevalence and underlying mechanism behind the response shift in pediatric population is still largely unknown. Research is needed that applies advanced quantitative and qualitative methods to address the puzzle. The last but not least pivotal issue is the urgent demand for developing methodology to monitor and compare PROs from childhood to adulthood, especially in children with chronic conditions (e.g., diabetes) or survivors of traumatic events (e.g., cancer, life-limiting conditions). Current PRO measurement systems are challenged by this issue because pediatric and adult instruments emphasize different developmental issues, and they were not calibrated on the same metric. We truly need developmental and life-course theories to generate a latent continuum of PROs (e.g., social functioning) that depicts different levels of functional status appropriately for children and adults. Such theories should be integrated with advanced psychometric methodology to link and calibrate PRO items and measurement tools across different developmental stages—from childhood to adulthood. Researchers of the National Institutes of Health sponsored

Patient Reported Outcomes Measurement Information System (PROMIS<sup>®</sup>) [19] are calibrating items in selected domains (i.e., pain behavior, physical functioning, etc.) across adult and pediatric populations. Although in early stages, this idea of a common metric for measuring PRO across age groups might be reaching fruition.

Children's health is our society's ultimate treasure. Instrument development for pediatric PROs is an ongoing endeavor. As PRO researchers, we are passionate and optimistic about the future, but we still have a long way to go.

## References

1. Simpson, J., & Speake, J. (2009). *The Oxford dictionary of proverbs*. New York, NY: Oxford University Press.
2. Riley, A. W. (2004). Evidence that school-age children can self-report on their health. *Ambulatory Pediatrics*, 4(4 Suppl), 371–376.
3. Fayed, N., de Camargo, O. K., Elahi, I., Dubey, A., Fernandes R. M., Houtrow, A., et al. (2013). Patient-important activity and participation outcomes in clinical trials involving children with chronic conditions. *Quality of Life Research*. doi:10.1007/s11136-013-0483-9.
4. Ahuja, B., Klassen, A. F., Satz, R., Malhotra, K., Tsangaris, E., Ventresca, M., et al. (2013). A review of patient-reported outcomes for children and adolescents with obesity. *Quality of Life Research*. doi:10.1007/s11136-013-0459-9.
5. Anthony, S. J., Selkirk, E., Sung, L., Klaassen, R. J., Dix, D., Scheinmann, K., et al. (2013). Considering quality of life for children with cancer: A systematic review of patient-reported outcome measures and the development of a conceptual model. *Quality of Life Research*. doi:10.1007/s11136-013-0482-x.
6. Ravens-Sieberer, U., Herdman, M., Devine, J., Otto, C., Bullinger, M., Rose, M., et al. (2013). The European KIDSCREEN approach to measure quality of life and well-being in children: Development, current application, and future advances. *Quality of Life Research*. doi:10.1007/s11136-013-0428-3.
7. Matza, L. S., Swensen, A. R., Flood, E. M., Secnik, K., & Leidy, N. K. (2004). Assessment of health-related quality of life in children: A review of conceptual, methodological, and regulatory issues. *Value in Health*, 7(1), 79–92.
8. Ravens-Sieberer, U., Erhart, M., Wille, N., Wetzel, R., Nickel, J., & Bullinger, M. (2006). Generic health-related quality-of-life assessment in children and adolescents: Methodological considerations. *Pharmacoeconomics*, 24(12), 1199–1220.
9. Solans, M., Pane, S., Estrada, M. D., Serra-Sutton, V., Berra, S., Herdman, M., et al. (2008). Health-related quality of life measurement in children and adolescents: A systematic review of generic and disease-specific instruments. *Value in Health*, 11(4), 742–764.
10. Kenzik, K., Huang, I. C., Tuli, S., Nackashi, J., Revicki, D., & Shenkman, E. (2011). Head-to-head comparisons of four legacy pediatric health-related quality of life instruments: A study on parent proxy-report. *Quality of Life Research*, 20(1 Suppl), 98.
11. Wilson, I. B., & Cleary, P. D. (1995). Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA*, 273(1), 59–65.
12. Nightingale, C. L., Quinn, G. P., Shenkman, E. A., Curbow, B. A., Zebrack, B. J., Krull, K. R., et al. (2011). Health-related quality of life of young adult survivors of childhood cancer: A

- review of qualitative studies. *Journal of Adolescent and Young Adult Oncology*, 1(3), 124–132.
13. Huang, I. C., Brinkman, T. M., Kenzik, K., Gurney, J. G., Ness, K. K., Lancot, J., et al. (2013). Association between the prevalence of symptoms and health-related quality of life in adult survivors of childhood cancer: A report from the St Jude Lifetime cohort study. *Journal of Clinical Oncology*, 31(33), 4242–4251.
  14. Varni, J. W., Seid, M., & Rode, C. A. (1999). The PedsQL: Measurement model for the pediatric quality of life inventory. *Medical Care*, 37(2), 126–139.
  15. Huang, I. C., Osakwe, O., Thompson, L., DeWalt, D., Revicki, D., & Shenkman, E. (2012). Pediatrician's capability, willingness, and barriers of measuring patient-reported outcomes in clinical practice: A national survey in the United States. *Quality of Life Research*, 21(1 Suppl), 50.
  16. Sheline, G. E. (1975). Radiation therapy of tumors of the central nervous system in childhood. *Cancer*, 35(3 Suppl), 957–964.
  17. Revicki, D., Hays, R. D., Cella, D., & Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology*, 61(2), 102–109.
  18. Schwartz, C. E., Feinberg, R. G., Jilinskaia, E., & Applegate, J. C. (1999). An evaluation of a psychosocial intervention for survivors of childhood cancer: Paradoxical effects of response shift over time. *Psycho-oncology*, 8(4), 344–354.
  19. Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., et al. (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.