

# Is There Hope for Juvenile Idiopathic Arthritis Patient-reported Outcomes in Support of Clinical Decisions?

Imagine if there were reliable, patient-reported assessments of disease activity that could be reported remotely to the physician, who could then make therapy decisions without needing to see the patient in person. While such a tool could greatly facilitate patient care, it is still just a fantasy.

In a study reported in this issue of *The Journal*, Dijkstra, *et al*<sup>1</sup> sought to replicate previous work with patient-reported outcomes (PRO) using a new format of the assessment tool. In particular they sought to validate a pictorial assessment aid for patients with juvenile idiopathic arthritis (JIA) to perform their own joint counts in the context of typical clinical care; and while the purpose was ostensibly to validate an instrument, they also report their results.

Patients at ages of those in this study (12-21 yrs) should be sufficiently mature to read and follow instructions and give answers at a level similar to adults. The patient assessments included a self-administered joint count based on a printed definition of active arthritis that included pain, swelling, and limited motion (but not chronically limited motion) in a joint, and a Child Health Assessment Questionnaire (CHAQ) that included visual analog scales (VAS) for well-being and pain. As well, the physician reported a joint count and VAS results for current disease activity. Joint counts by patients and by physician were repeated at subsequent visits, and apparently the CHAQ and physician VAS for current disease activity were repeated as well; however, only results from the first and second visits were reported. The physician-assessed joint count (by one physician) was taken as the reference in the statistical comparisons because no clinically practical gold standard has yet been established.

The results seem similar to studies on juvenile joint counts and are comparable to studies in adults. On the one hand, it is reassuring that these PRO may conceivably be used in patients as young as 12 years of age; on the other hand, patient agreement with physician joint count is generally still disappointingly low. There may be a variety of reasons for the difference between physician and patient joint count, not the least of which is that physician assessment may not be an ideal reference group<sup>2,3</sup>. A recent study claimed that, in their JIA cases involving knees, more

than one-third of patients clinically assessed as having inactive disease had subclinical activity detectable by contrast-enhanced magnetic resonance imaging (MRI), while nearly half of patients with JIA whose disease was “considered clinically active showed no signs of MRI-based synovitis.”<sup>3</sup> Although one could debate whether it is MRI or physician assessment that is unreliable, the study underscores the differences that different methods produce.

Most discrepancies between patient and physician assessments fell between adjacent categories, which may mean that differences are not large enough to alter therapy decisions. But Dijkstra, *et al* did not examine whether these differences in PRO might lead to alternative therapy decisions.

Having only one person do assessments seems to be a favorite method among investigators; this usually reduces variability and potentially increases sensitivity to detect change. However, this approach is not ideal for assessing agreement when there is no gold standard. As mentioned in their discussion<sup>1</sup>, the proper way is to increase the number of evaluating physicians. While inconvenient or worse in the everyday clinical setting, multiple assessment by different physicians would be recommended when doing studies to determine whether patient assessment could be used as adjunct to or to replace physician assessment. The book by Fleiss, *et al* includes a good discussion of measuring agreement<sup>4</sup>. One of several possible schemes available involves pairwise assessments of patients using different pairings of several physicians, so that every possible pairing is repeated at least once over time.

Intraclass correlation coefficient (ICC) is another issue that arises in translating results from one population to another: ICC does not generalize well from one population to another because some variance components used in the formula depend on the actual patient mix. A personal recommendation is to report the components of variance that are used to compute the ICC, as well as the ICC themselves and the formulae or descriptions of the models used to compute them. This would provide a way to anti-

---

See PRO in JIA, page 527

pate what the ICC might look like in another population. (This is of statistical concern, but statisticians are involved in every study and they also have to read these articles; why not provide some information useful to them?)

In agreement with the conclusion of Dijkstra, *et al*, reasonable next steps would be to investigate whether, with instruction, patient and physician assessments may converge, and whether PRO are able to detect clinically useful changes over time. Some of these questions have been studied in adult populations<sup>5</sup>; what is not known is how well or down to what age group results may be applied to younger patients.

#### ACKNOWLEDGMENT

The author thanks Randy Lehmer, MD, for his clinical practice insights.

**RAOUL J. BURCHETTE**, MS,  
Research and Evaluation Department,  
Kaiser Permanente Southern California,  
Pasadena, California, USA

Address correspondence to R.J. Burchette, 100 S. Los Robles Ave.,  
2nd Floor, Pasadena, California 91101, USA.  
E-mail: Raoul.J.Burchette@kp.org

#### REFERENCES

1. Dijkstra ME, Anink J, van Pelt PA, Hazes JM, van Suijlekom-Smit LW. Patient-reported joint count in juvenile idiopathic arthritis: the reliability of a manikin format. *J Rheumatol* 2015;42:527-33.
2. Guzmán J, Burgos-Vargas R, Duarte-Salazar C, Gómez-Mora P. Reliability of the articular examination in children with juvenile rheumatoid arthritis: interobserver agreement and sources of disagreement. *J Rheumatol* 1995;22:2331-6.
3. Hemke R, Maas M, van Veenendaal M, Dolman KM, van Rossum MA, van den Berg JM, et al. Contrast-enhanced MRI compared with the physical examination in the evaluation of disease activity in juvenile idiopathic arthritis. *Eur Radiol* 2014;24:327-34.
4. Fleiss JL, Levin B, Paik MC. The measurement of interrater agreement. In: *Statistical methods for rates and proportions*. 3rd ed. Hoboken: John Wiley; 2003:598-626.
5. Cheung PP, Lahiri M, Teng GG, Lim AY, Lau TC, Lateef A, et al. A randomized controlled trial for improving patient self-assessment of synovitis in rheumatoid arthritis with education by ultrasonography: the RAEUS Study. *Rheumatology* 2014; Dec 3 (E-pub ahead of print).

*J Rheumatol* 2015;42:358-9; doi:10.3899/jrheum.141621