can overestimate the scores on performance-based tests. Many clinicians and investigators would say that total hip and knee arthroplasty are 2 of the most successful surgeries in the field of orthopedics. Much of the justification for such labeling stems from self-reported instruments that can be heavily swayed by the powerful capacity for arthroplasty to reduce pain. These 2 surgeries may represent some of the most extreme and consistent reductions in pain in today’s health care environment.

Physical therapy is well positioned as a profession to address the common and often lingering impairments in musculoskeletal performance that have consistently been related to residual locomotor deficits. Clinicians may miss opportunities to maximize their patients’ postoperative gains if they fail to take a more comprehensive view of physical function by adding performance-based measures to their outcomes array. I look forward to reading future work on this topic from Stratford and colleagues as they continue to pursue this line of research.

References
1 Stratford PW, Kennedy DM, Maly MR, MacIntyre NJ. Quantifying self-report measures’ overestimation of mobility scores postarthroplasty. Phys Ther. 2010;90:1288–1296.
2 Roadmap for Implementing Value Driven Healthcare in the Traditional Medicare Fee-for-Service Program. Centers for Medicare and Medicaid Services Web site. Available at: gov/QualityInitiative-GenInfo/downloads/VBPRoadmap_OEA_1-16_508.pdf-20100402. Accessed May 31, 2010.
3 Parent E, Moffet H. Comparative responsiveness of locomotor tests and questionnaires used to follow early recovery after total knee arthroplasty. Arch Phys Med Rehabil. 2002;83:70–80.
4 Ouellet D, Moffet H. Locomotor deficits before and two months after knee arthroplasty. Arthritis Rheum. 2002;47:484–493.
5 Mizner RL, Petterson SC, Snyder-Mackler L. Quadriceps strength and the time course of functional recovery after total knee arthroplasty. J Orthop Sports Phys Ther. 2005;35:424–436.
6 Petterson SC, Mizner RL, Stevens JE, et al. Improved function from progressive strengthening interventions after total knee arthroplasty: a randomized clinical trial with an imbedded prospective cohort. Arthritis Rheum. 2009;61:174–185.
7 Stratford PW, Kennedy DM. Performance measures were necessary to obtain a complete picture of osteoarthritic patients. J Clin Epidemiol. 2006;59:160–167.
8 Valtonen A, Pöyhönen T, Heinonen A, Sipilä S. Muscle deficits persist after unilateral knee replacement and have implications for rehabilitation. Phys Ther. 2009;89:1072–1079.
9 Kowalchuk DA, Harner CD, Fu FH, et al. Prediction of patient-reported outcome after single-bundle anterior cruciate ligament reconstruction. Arthroscopy. 2009;25:457–463.
10 Kennedy DM, Stratford PW, Riddle DL, et al. Assessing recovery and establishing prognosis following total knee arthroplasty. Phys Ther. 2008;88:22–32.
11 Ferrara PE, Rabini A, Maggi L, et al. Effect of pre-operative physiotherapy in patients with end-stage osteoarthritis undergoing hip arthroplasty. Clin Rehabil. 2008;22:977–986.

Invited Commentary

Yong Hao Pua, Kim L. Bennell

Much has been written in the literature about the close association between self-report measures of physical function and pain in patients with upper-extremity1,2 and lower-extremity3–6 musculoskeletal conditions. That pain wields a strong influence on self-report measures of physical function is neither a trivial nor a purely academic finding: there are instances where pain and physical function can uncouple, which, in turn, limits the ability of self-report measures to assess functional ability. Specifically, in the context of patients who have recently undergone a knee or hip arthroplasty, at least 2 studies5,6 have shown that these patients “felt better but did worse”: that is, self-report measures of pain and physical function improved significantly, but walking and stair-climbing performance was no better—or even worse—than the prearthroplasty performance.

Against this background, the article by Stratford and colleagues7 provides thought-provoking data on the extent to which self-report measures of function—namely, the Western Ontario and McMaster Universities Osteoarthritis Index physical function subscale (WOMAC-PF) and the Lower Extremity Functional Scale (LEFS)—systematically overestimated the patients’ performance on the Six-Minute Walk Test (6MWT) and the Timed “Up & Go” Test (TUG). Stratford and colleagues suggest that their point estimates of fixed bias could help readers reinterpret findings of previous studies that examined only self-report measures. Stratford and colleagues’ work is an important contribution, and the aim of our commentary is to highlight 2 aspects of their study that deserve special consideration.

First, the study by Stratford et al7 examined the 6MWT and the TUG, and it is unknown whether the observed fixed bias would remain invariant across other performance measures such as stair-climbing and gait speed tests. Yet, it may be necessary to extend the regression analyses of Stratford et al7 to these measures because our previous work4...
and that of Stratford et al⁸ suggest that the various performance measures were differentially correlated with self-report measures of pain and physical function in patients with hip or knee osteoarthritis. Accordingly, to the extent that the fixed biases reported by Stratford et al⁷ remained stable across the different performance measures, greater confidence may be placed on the robustness and generalizability of the data.

A second issue relates to the conservativeness of the criteria used by Stratford et al⁷ to ascertain whether the observed fixed biases were clinically important. In the first place, it has been argued that the minimal important difference (also known as the minimal clinically important difference) for a group of patients is lower than that for an individual patient,⁹ whereas the minimal important difference (based on a small change) for an individual patient is reportedly lower than the minimal detectable change.¹⁰–¹³ Stratford and colleagues’ criteria were 9 points for both the WOMAC-PF and the LEFS, which represented the minimal detectable change for the individual patient.⁴ Accordingly, it is likely that Stratford and colleagues’ yardstick was too stringent when appraising within-group differences, and, therefore, the authors may have underestimated the significance of their findings. Finally, so that readers can better appreciate the clinical significance of Stratford and colleagues’ findings, it may be useful for the authors to report the proportion of patients whose systematic biases exceeded the minimal important difference⁴¹⁴,¹⁵

To sum up, the article by Stratford et al⁷ reminds us about the substantial mismatch between self-report and performance measures of physical function in patients who recently underwent a knee or hip arthroplasty. Although Stratford et al aim to provide readers with useful data to appraise past research work, they also should be thanked for providing a compelling argument for the inclusion of performance measures of physical function in present and future research efforts so that we can draw a more comprehensive picture of the patient’s lower-extremity functional status.

Y.H. Pua, PhD, is Physiotherapist, Department of Physiotherapy, Singapore General Hospital, Singapore. Address all correspondence to Dr Pua at: pua.yong.hao@sgh.com.sg.

K.B. Bennell, PhD, is Professor, Centre for Health, Exercise and Sports Medicine, Department of Physiotherapy, School of Health Sciences, The University of Melbourne, Melbourne, Victoria, Australia.

DOI: 10.2522/ptj.20100058.ic2

References

1 Midha R, Noble J, Patel V, Ho PH, et al. Prospective analysis of relationships of outcome measures for ulnar neuropathy at the elbow. Can J Neurol Sci. 2001;28: 239–244.
2 Doornberg JN, Ring D, Fabian LM, et al. Pain dominates measurements of elbow function and health status. J Bone Joint Surg Am. 2005;87:1725–1731.
3 Terwee CB, van der Slkke RM, van Lummel RC, et al. Self-reported physical functioning was more influenced by pain than performance-based physical functioning in knee-osteoarthritis patients. J Clin Epidemiol. 2006;59:724–731.
4 Pua YH, Cowan SM, Wrigley TV, Bennell KL. The Lower Extremity Functional Scale could be an alternative to the Western Ontario and McMaster Universities Osteoarthritis Index physical function scale. J Clin Epidemiol. 2009;62:1103–1111.

5 Stratford PW, Kennedy DM. Performance measures were necessary to obtain a complete picture of osteoarthritic patients. J Clin Epidemiol. 2006;59:160–167.
6 Parent E, Moffet H. Comparative responsiveness of locomotor tests and questionnaires used to follow early recovery after total knee arthroplasty. Arch Phys Med Rehabil. 2002;83:70–80.
7 Stratford PW, Kennedy DM, Maly MR, MacIntyre NJ. Quantifying self-report measures’ overestimation of mobility scores postarthroplasty Phys Ther. 2010;90: 1288–1296.
8 Stratford PW, Kennedy DM, Woodhouse LJ. Performance measures provide assessments of pain and function in people with advanced osteoarthritis of the hip or knee. Phys Ther. 2006;86:1489–1496.
9 Goldsmith CH, Boers M, Bombardier C, Tugwell P. OMERACT Committee. Criteria for clinically important changes in outcomes: development, scoring and evaluation of rheumatoid arthritis patient and trial profiles. J Rheumatol. 1993;20: 561–565.
10 Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care. 2003;41:582–592.
11 Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intrapatient changes in health-related quality of life. J Clin Epidemiol. 1999;52: 861–873.
12 Turner D, Schunemann HJ, Griffith LE et al. The minimal detectable change cannot reliably replace the minimal important difference. J Clin Epidemiol. 2010;63: 28–36.
13 Rejas J, Pardo A, Ruiz MA. Standard error of measurement as a valid alternative to minimally important difference for evaluating the magnitude of changes in patient-reported outcomes measures. J Clin Epidemiol. 2008;61:350–356.
14 Schmit JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. J Clin Epidemiol. 2004;57: 1008–1018.
15 Guyatt G, Schunemann H. How can quality of life researchers make their work more useful to health workers and their patients? Qual Life Res. 2007;16: 1097–1105.