Minimal Important Difference and Responsiveness to Change of the SF-36 in Patients with Psoriatic Arthritis Receiving Tumor Necrosis Factor-α Blockers

To the Editor:

The Medical Outcome Health Survey Short-form 36 (SF-36) has been validated in psoriatic arthritis (PsA)[1] and its psychometric assumptions tested[2,3], but no data on its minimal important difference (MID) exist. MID is the smallest change in a patient-reported outcome (PRO) score that patients perceive as a meaningful change[4]. It differentiates treatment response and may serve as threshold for therapy change. We examineed the MID and responsiveness to change of SF-36 scales in patients with psoriatic arthritis (PsA) undergoing treatment with tumor necrosis factor-α (TNF-α) blockers.

Twenty consecutive patients with active PsA fulfilling the CASPAR criteria[5] were recruited to receive TNF-α blockers. After 12 weeks, 9 patients continued TNF-α blockers (Group 1); 11 patients discontinued due to financial constraints. PRO including pain, patient's global health assessment (PGA), Health Assessment Questionnaire (HAQ)[6], and SF-36[7] were collected at baseline and Weeks 12, 24, 36, and 52. At each visit, patients answered an anchor question on general health status, "How do you rate your current health status as compared to last visit: (1) much better, (2) slightly better, (3) similar, (4) slightly worse, and (5) much worse." The differences in PRO scores between 2 consecutive visits (e.g., Week 12 – Week 0, Week 24 – Week 12) were calculated. In 4 followups, 78 sets of PRO scores were analyzed. MID estimates were calculated as the mean change of PRO variables in those who rated their disease as "slightly better" or "slightly worse" in the anchor question, which represented a minimal change that is perceived by patients as relevant[8]. The change in anchor and the change in PRO scores should have a correlation coefficient of at least 0.3[9]. The responsiveness of PRO were compared using the effect size (ES) and the standardized response mean (SRM) in patients who rated "slightly better" and "better" in the anchor question. The ES is small if < 0.2, medium if 0.3–0.5, and large if > 0.5[10]. Exploratory analyses stratified by sex and severity of disease (HAQ > 1.0) were performed on the MID for various PRO.

Table 1 shows the demographic characteristics of patients at baseline and Week 12 and 52. PRO in both groups improved to a similar extent at Week 12, but worsened in Group 2 after Week 12. We found that 34.6%, 21.3%, 10.3%, 26.9%, and 6.4% of patients rated their general health status as "much better," "slightly better," "similar," "slightly worse," and "much worse" compared to last visit in the anchor question. The changes in SF-36 scales and PRO between the "slightly better" and "slightly worse" groups were significantly different (Table 2). Pain, PGA, HAQ, and the SF-36 scales for physical function (PF), bodily pain (BP), general health (GH), and physical component summary score (PCS) had the desirable Spearman's rho of > 0.3. The correlations were less for VT, role emotional (RE), mental health (MH), and mental component summary score (MCS). The ES and SRM for PF, BP, GH, and PCS were moderate (range 0.35 to 0.59). In the exploratory analyses, no difference in MID values was noted when stratified by sex, except in the social functioning (SF) scale. The MID estimates for improvement were larger in males (data not shown). There was no difference in MID when stratifying by HAQ > 1.0.

MID has been defined as "the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side-effects and excessive cost, a change in the patient's management"[4]. Determining MID is especially important because statistical significance is not equivalent to clinical significance. A change of therapy may be necessary when patients are not improving to a meaningful extent. Our study is the first to determine MID and responsiveness for SF-36 in patients with PsA as classified by the CASPAR criteria undergoing TNF-α blocker therapy. This provided additional evidence of the validity of the SF-36 in PsA. Improvement after TNF-α blockers and deterioration on stopping were expected and therefore

Table 1. Demographic characteristics of PsA patients at baseline, Week 12, and Week 52. Group 1 continued TNF blockers after 12 weeks. Group 2 discontinued TNF blockers after 12 weeks.

Characteristic	Baseline		Week 12		Week 52	
	Group 1 (n = 9)	Group 2 (n = 11)	Group 1 (n = 9)	Group 2 (n = 11)	Group 1 (n = 9)	Group 2 (n = 11)
Age, yrs	47.8 (13.0)	51.5 (11.1)				
Illness duration, yrs	5.7 (5.7)	9.5 (7.5)				
Sex, % male	54	63.6				
Tender joint (0–68)	11.2 (6.0)	9.8 (4.0)	3.0 (6.1)	3.4 (4.4)	2.0 (3.5)	7.0 (10.2)
Swollen joint (0–66)	4.7 (1.6)	2.9 (2.5)	1.2 (1.9)	1.1 (1.4)	0	2.2 (0.7)**
DAS28	6.14 (0.83)	5.55 (0.67)	2.70 (1.37)	2.83 (1.28)	1.24 (0.44)	1.38 (0.44)**
Pain (0–100)	67.8 (23.3)	62.7 (16.8)	14.4 (10.1)	33.2 (23.5)*	20.6 (20.8)	49.5 (26.5)*
PGA (0–100)	67.8 (20.5)	58.2 (16.6)	28.9 (22.6)	32.3 (22.1)	22.3 (21.3)	42.5 (25.5)
HAQ	1.16 (0.59)	1.02 (0.68)	0.43 (0.53)	0.80 (0.74)	0.50 (0.18)	0.86 (0.27)
SF-36 scales						
PF	51.5 (23.5)	57.7 (21.8)	76.7 (25.0)	73.3 (22.1)	70.6 (30.4)	54.5 (28.0)
RP	19.4 (34.9)	20.5 (40.0)	75.0 (35.6)	69.4 (42.9)	59.3 (42.1)	25.0 (43.3)
BP	30.7 (12.5)	42.8 (21.4)	71.6 (20.1)	54.0 (17.1)	67.8 (22.4)	40.2 (19.2)**
GH	23.8 (8.0)	40.5 (11.0)**	49.6 (20.4)	47.1 (13.0)	44.4 (25.2)	27.5 (10.6)
VT	43.9 (9.3)	41.8 (11.2)	55.6 (21.0)	45.6 (15.3)	52.8 (21.4)	36.4 (19.0)
SF	50.1 (29.2)	55.9 (31.3)	77.8 (21.4)	77.8 (22.3)	75.0 (18.8)	52.2 (27.8)
RE	29.6 (42.3)	24.2 (39.7)	81.5 (33.8)	74.1 (43.4)	62.5 (37.5)	21.2 (40.2)*
MH	50.7 (23.1)	57.5 (11.5)	60.0 (20.9)	68.0 (11.1)	61.5 (20.7)	53.2 (15.1)
PCS	22.3 (7.6)	28.0 (9.3)	41.9 (13.3)	36.6 (10.5)	41.2 (16.4)	26.2 (13.9)
MCS	38.8 (12.6)	38.8 (9.2)	48.0 (13.3)	49.2 (9.9)	44.8 (13.4)	37.2 (11.0)

DAS28: Disease Activity Score (28 joints); PGA: Patients assessment of global health; HAQ: Health Assessment Questionnaire; PF: physical functioning; RP: role limitation due to physical problem; BP: bodily pain; GH: general health; VT: vitality; SF: social functioning; RE: role limitation due to emotional problem; MH: mental health; PCS: physical component summary score; MCS: mental component score. * If p < 0.05. ** If p < 0.01.

Table 2. Responsiveness and minimal important difference (MID) estimates for SF-36 and patient-reported outcomes.

	MID Improvement, (n = 17)	MID Deterioration, (n = 21)	Correlation with Anchor	Effect Size	SRM
Pain	–14.71 (31.25)	–0.95 (18.82)**	0.30*	–0.55	–0.49
PGA	–11.76 (25.06)	–2.86 (18.88)*	0.31**	–0.50	–0.55
HAQ	–0.27 (0.42)	0.095 (0.18)**	0.30**	–0.22	–0.22
SF-36 scales					
PF	4.41 (14.99)	–6.25 (18.77)**	–0.34**	0.35	0.37
RP	11.76 (44.30)	–11.25 (36.70)*	–0.29*	0.41	0.39
BP	12.35 (19.83)	–5.45 (15.63)**	–0.41**	0.53	0.59
GH	–2.94 (12.34)	–3.75 (15.02)*	–0.30**	0.25	0.24
VT	7.94 (11.46)	–5.25 (15.68)**	–0.28**	0.28	0.35
SF	5.06 (22.50)	–2.78 (17.45)*	–0.27*	0.26	0.31
RE	3.96 (52.60)	–10.0 (37.62)*	–0.23*	0.27	0.26
MH	1.41 (8.36)	0.00 (14.85)*	–0.28*	0.19	0.31
PCS	3.74 (8.51)	–3.97 (10.46)**	–0.43**	0.49	0.55
MCS	1.77 (8.60)	–0.70 (8.37)*	–0.24*	0.19	0.28

* If p < 0.05. ** If p < 0.01. Definitions as in Table 1.

we observed the MID estimates on both ends. Desirable correlations were observed between the anchor health status and PF, BP, GH, PCS, and HAQ. Although the correlations between the anchor question and some PRO variables were < 0.3, they were still significantly associated with the outcome of interest. This illustrates that changes in pain, PGA, and physical function may be more accurate at detecting small changes in the patient's perceived status. Vitality and mental health may be affected by other factors not exclusively related to PsA activity. Fatigue (the opposite of vitality) was shown to vary with time and was associated with other factors like sex, psychological distress, and medical comorbidities in PsA[11,12]. The SF-36 MCS was also weaker in distinguishing drug or placebo effect in a phase III drug trial[13]. The MID scores were asymmetrical in worsening and improvement for pain, PGA, and HAQ. These asymmetrical changes in MID have been reported in other studies[14,15]. There is a smaller change in scores for patients to perceive deterioration than for improvement, which may be related to the high expectation of treatment effect from patients undergoing a drug trial.

For the responsiveness analyses, moderate ES and SRM were seen for SF-36 BP and PCS. Comparing measurements for physical function, HAQ, and SF-36 PCS, the latter had better ES and SRM. The SF-36 was also reported by Husted, *et al* to be more responsive as compared to HAQ and the Arthritis Impact Measurement Scale[16].

There are several limitations to our study. The small sample size yielded large SD in the MID estimates and introduced unreliability to a certain extent. The results were subject to recall bias when patients were asked about their change compared to last visit[8,9]. MID results may change when a different anchor is used. For example, participants may have selected "slightly better" even if they experienced a substantial change because they were unwilling to call their change "much better." Although this health status anchor question has not been validated in PsA, similar general health status anchors have been employed in studies in other connective tissue diseases[15,17,18,19]. Further, MID may change with baseline disease severity[15] and in different clinical settings[17]. Our study was performed in Han Chinese patients with long disease duration, from a tertiary referral center, and with severe peripheral arthritis. These results may not be generalizable to PsA patients from other ethnic groups with milder disease severity, shorter disease duration, or predominant axial manifestation. Despite all these limitations, the MID estimates for pain, PGA, and HAQ were similar to those of other studies[15,17,18,19,20]. The SF-36 PCS, pain, and PGA proved responsive to short-term changes in PsA.

YING-YING LEUNG, MD, Department of Medicine and Geriatrics, Tai Po Hospital; TRACY YANER ZHU, MD; LAI SHAN TAM, MD, Department of Medicine and Therapeutics, Chinese University of Hong Kong; EMILY WAI-LIN KUN, MD, Department of Medicine and Geriatrics, Tai Po Hospital; EDMUND KWOK-MING LI, MD, Department of Medicine and Therapeutics, Chinese University of Hong Kong, Hong Kong, China. Address correspondence to Dr. Y-Y. Leung, Department of Medicine and Geriatrics, Tai Po Hospital, HKSAR, 9 Chuen On Road, Tai Po, NT, Hong Kong. E-mail: katyccc@hotmail.com

REFERENCES
1. Husted JA, Gladman DD, Farewell VT, Long JA, Cook RJ. Validating the SF-36 health survey questionnaire in patients with psoriatic arthritis. J Rheumatol 1997;24:511-7.
2. Leung YY, Ho KW, Zhu TY, Tam LS, Kun EW, Li EK. Testing scaling assumptions, reliability and validity of Medical Outcomes Study Short-form 36 health survey in psoriatic arthritis. Rheumatology 2010;49:1495-501.
3. Leung YY, Tam LS, Kun EW, Ho KW, Li EK. Comparison of 4 functional indexes in psoriatic arthritis with axial or peripheral disease subgroups using Rasch analyses. J Rheumatol 2008;35:1613-21.
4. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials 1989;10:407-15.
5. Taylor W, Gladman D, Helliwell P, Marchesoni A, Mease P, Meilants H, CASPAR Study Group. Classification criteria for psoriatic arthritis: Development of new criteria from a large international study. Arthritis Rheum 2006;54:2665-73.
6. Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the Health Assessment Questionnaire, disability and pain scales. J Rheumatol 1982;9:789-93.
7. Lam CL, Tse EY, Gandek B, Fong DY. The SF-36 summary scales were valid, reliable, and equivalent in a Chinese population. J Clin Epidemiol 2005;58:815-22.
8. De Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL,

Bouter LM. Minimal changes in health status questionnaires: Distinction between minimally detectable change and minimally important change. Health Qual Life Outcomes 2006;4:54.

9. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61:102-9.

10. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates; 1998.

11. Husted JA, Tom BD, Schentag CT, Farewell VT, Gladman DD. Occurrence and correlates of fatigue in psoriatic arthritis. Ann Rheum Dis 2009;68:1553-8.

12. Husted JA, Tom BD, Farewell VT, Gladman DD. Longitudinal analysis of fatigue in psoriatic arthritis. J Rheumatol 2010;37:1878-84.

13. Gladman DD, Tom BD, Mease PJ, Farewell VT. Informing response criteria for psoriatic arthritis. I: Discrimination models based on data from 3 anti-tumor necrosis factor randomized studies. J Rheumatol 2010;37:1892-7.

14. Ware JE, Snow K, Kosinski M, Gandek B. SF-36 Health survey: Manual and interpretation guide. Boston: The Health Institute; 1993.

15. Wheaton L, Pope J. The minimally important difference for patient-reported outcomes in spondyloarthropathies including pain, fatigue, sleep, and Health Assessment Questionnaire. J Rheumatol 2010;37:816-22.

16. Husted JA, Gladman DD, Cook RJ, Farewell VT. Responsiveness of health status instruments to changes in articular status and perceived health in patients with psoriatic arthritis. J Rheumatol 1998;25:2146-55.

17. Pope JE, Khanna D, Norrie D, Ouimet J. The minimally important difference for the Health Assessment Questionnaire in rheumatoid arthritis clinical practice is smaller than in randomized controlled trials. J Rheumatol 2009;36:254-9.

18. Colangelo KJ, Pope JE, Peschken C. The minimally important difference for patient reported outcomes in systemic lupus erythematosus including the HAQ-DI, pain, fatigue, and SF-36. J Rheumatol 2009;36:2231-7.

19. Kwok T, Pope JE. Minimally important difference for patient-reported outcomes in psoriatic arthritis: Health Assessment Questionnaire and pain, fatigue, and global visual analog scales. J Rheumatol 2010;37:1024-8.

20. Mease P, Ganguly L, Wanke E, Yu E, Singh A. How much improvement in functional status is considered important by patients with active psoriatic arthritis: applying the outcome measures in rheumatoid arthritis clinical trials. Ann Rheum Dis 2004;63 Suppl 1:391.