Minimal Clinically Important Difference, Low Disease Activity State, and Patient Acceptable Symptom State: Methodological Issues

FLORENCE TUBACH, GEORGE A. WELLS, PHILIPPE RAVAUD, and MAXIME DOUGADOS

ABSTRACT. The importance of determining a minimal clinically important difference (MCID) and a low disease activity state (LDAS) as treatment targets in clinical trials no longer needs to be demonstrated. However, many methodological issues remain: whether these thresholds should be defined for each criterion or for composite criteria, whether there is a difference between the LDAS and patient acceptable symptom state (PASS), how to determine these thresholds (i.e., the wording of the questions and the statistical approach), and whether there are confounding factors in their evaluation. We consider these methodological issues and discuss their impact. Methods to determine the thresholds must be standardized, and recommendations could be endorsed by an OMERACT module. Threshold values for the MCID and LDAS should be determined according to data-driven and experts' opinions and approaches. (J Rheumatol 2005;32:2025–9)

Key Indexing Terms:

MINIMAL CLINICALLY IMPORTANT DIFFERENCE DISEASE ACTIVITY
LOW DISEASE ACTIVITY STATE OUTCOME ASSESSMENT PATIENT PERSPECTIVE

Introduction

In clinical trials, outcome criteria are usually measured over time. When health status is measured using a continuous scale, results are presented at the group level (mean difference between baseline and final visits); thus, readers need some information about the clinical relevance of the observed results. They need to know whether an observed difference constitutes a trivial or an important improvement in symptoms [minimum clinically important improvement (MCII)[1]] or whether the observed change leads to an acceptable state according to the patient and/or physician [low disease activity state (LDAS)[2] or patient acceptable symptom state (PASS)[3]]. The importance of determining a minimal clinically important difference and a low disease activity state as treatment targets in clinical trials no longer needs to be demonstrated[2,4]. However, many methodological issues remain: whether these thresholds should be defined by use of a direct or an indirect approach, whether there is a difference between PASS and LDAS, how to determine these thresholds (i.e., the wording of questions and the statistical approach), and whether there are confounding factors for their evaluation.

From the Département d'Epidémiologie, Biostatistique et Recherche Clinique, Groupe Hospitalier Bichat-Claude Bernard (AP–HP), INSERM U738, Faculté Xavier Bichat (Université Paris 7), Paris, France; Department of Clinical Epidemiology and Community Medicine, University of Ottawa, Ottawa, Canada; and Service de Rhumatologie B, Hôpital Cochin (AP–HP), Paris, France.

F. Tubach, MD, INSERM U738, Département d'Epidémiologie, Biostatistique et Recherche Clinique, Groupe Hospitalier Bichat-Claude Bernard, Faculté Xavier Bichat; G.A. Wells, PhD, Department of Clinical Epidemiology and Community Medicine, University of Ottawa; P. Ravaud, MD, PhD, INSERM U738, Département d'Epidémiologie, Biostatistique et Recherche Clinique, Groupe Hospitalier Bichat-Claude Bernard, Faculté Xavier Bichat; M. Dougados, MD, Service de Rhumatologie B, Hôpital Cochin.

Address reprint requests to Dr. F. Tubach, Département d'Epidémiologie, Biostatistique et Recherche Clinique, INSERM U738, Hôpital Bichat, 46 rue Henri Huchard, 75018 Paris, France.

Determining MCII and LDAS/PASS for Each Criterion and Then Combined or for Composite Criteria

Because the concepts of improvement and an acceptable state (or LDAS) reflect how the patient feels in general, it is important to address all features of disease activity. Usually, health status is measured using several tools covering several dimensions (e.g., rheumatoid disorders: pain, functional impairment, patient global assessment, number of involved joints, and erythrocyte sedimentation rate). Kirwan[5] has shown that for many rheumatologists, no single variable could change sufficiently to signify a clinically important difference by itself; rather, combined, modest changes in a number of variables may be considered useful. In previous studies, relevant improvement was based on assessment of a global core set of different outcome criteria[6,7], using an indirect approach[8] (e.g., definitions obtained from participants' assessments of patient profiles, i.e., for measurements of relevant features). In a study of hip and knee osteoarthritis[3,9] the MCII and PASS were determined independently for each patient-related outcome, by use of the direct opinion-based approach[8] (i.e., from participants asked to define MCII and PASS for each of the measures relevant to disease activity). Thus, patients can be classified as being improved or not (or

achieving the PASS or not) in one criterion. To define a relevant clinical response, a group of experts could define improvement (or acceptable state) as satisfying the threshold in a fixed number of criteria (for instance, 2 out of 3).

Are Low Disease Activity State and Patient Acceptable Symptom State Different?

During OMERACT 6, LDAS was defined as an intermediate state between high disease activity and remission that could also be called "partial remission," and which is a useful treatment target for both physicians and patients. During OMERACT 7, the operational definition of LDAS in rheumatoid arthritis was derived by presenting profiles of patients with rheumatoid arthritis to 40 experts[10] instructed to consider that each profile corresponded to a patient starting methotrexate whose dosage had been increased to the expert's usual dosage. The profile described actual levels of measures relating to each feature of disease activity after at least 6 months of therapy at that dose. For each profile, experts were asked whether the patient was in a LDAS (i.e., whether the expert would keep the patient on methotrexate at that dosage). If experts thought the therapy had to be changed, patients were considered not to be in a LDAS.

The definition of LDAS is therefore closely linked to treatment decision-making and is based solely on the clinical experience of the physician. The advantage of this definition is that it is linked to an identifiable and relevant cutpoint in the global care of patients. The drawback is that the value of LDAS should be regularly updated as treatment options and knowledge evolve. The target will become outdated once future therapy allows for LDAS with a similar or lower toxicity level[2].

The PASS, however, has been defined as the value beyond which patients can consider themselves well[3]. The concept is not linked to treatment decisions and only addresses the symptomatic state; thus, it is based solely on the patient's perspective. Because the definitions of LDAS and PASS differ, the thresholds may differ. In a previous study[11], 177 outpatients in rheumatology [all with chronic rheumatic disease and familiar with assessing their pain according to a visual analog scale (VAS; 0 to 100 mm)] were asked the following: (1) "What is the level of pain above which you experience difficulties?" (This could be considered close to the external anchor for the PASS.) (2) "What is the level of pain above which you would consider taking a pain killer drug?" (This could be considered close to the external anchor for the LDAS.) The mean (± SD) pain level above which patients experienced difficulties was 38.8 ± 17.1 mm, and the mean pain level above which they would consider taking a pain killer was 50.4 ± 9.5 mm.

From the patients' perspective, therefore, LDAS and PASS seem different. They are 2 different intermediate states between high activity and remission. The first state addresses a symptomatic state above which a therapeutic decision should be taken and the second a relevant and desirable symptomatic state. Thus, the LDAS could be used as an inclusion criterion in trials and the PASS an outcome criterion.

Does the Wording of the External Anchor and the Response Modalities Affect the Results? How Should Questions About the External Anchor Be Asked?

In a cohort study of 1362 outpatients with knee or hip osteoarthritis (OA) requiring treatment with a nonsteroidal antiinflammatory drug (NSAID)[9], patients assessed their status regarding their OA at baseline and final visits (4 weeks later) by completing a pain and global assessment as measured on a VAS and the Western Ontario and MacMaster Universities (WOMAC) physical function subscale. At the final visit, patients were asked to assess the following:

1. *Patients' response to therapy*: "How would you rate your response to the NSAID medication you have received for your arthritis for 4 weeks?", measured on a 5-point Likert scale [(1A) two-thirds of patients] ranging from none: no good at all, ineffective drug; poor: some effect but unsatisfactory; fair: reasonable effect, but could be better; good: satisfactory effect with occasional episodes of pain or stiffness; to excellent: ideal response, virtually pain-free; and on a 15-point Likert scale [(1B) one-third of patients) ranging from –7, a very great deal worse, to +7, a very great deal better.
2. *Patients' opinion of their improvement:* "In your opinion, has the treatment received during the last 4 weeks notably improved your condition?", with a dichotomous response of yes or no.

Table 1 shows that the choice of the wording of the question (1 vs 2) and the response modalities (1A vs 1B) was arbitrary and distorted the results, as did the threshold chosen to define patients with an important improvement. The group of patients in whom MCII is determined and the wording of the items in the questionnaire to assess response to therapy should be chosen with the help of experts and be identical in all studies attempting to determine thresholds for MCII or LDAS/PASS.

After working on the wording of questions to determine MCII and PASS, a group of experts proposed to assess MCII and PASS based on questions like those in Figure 1, in which pain is used as the outcome criterion.

What Is the Impact of the Choice of Statistical Approach?

As concluded in the OMERACT 6 module on the minimal clinically important difference (MCID)[1], in determining an MCID for an outcome measure, 3 components are needed: an indicator that change has occurred or that a difference exists, a valid assignment of the importance of the change, and an appropriate method to determine the threshold within the distribution of changes. To determine the MCII and the LDAS/PASS, patient global ratings as described by

Table 1. Influence of response modalities and of wording of questions related to the external anchor (real improvement) for determining the minimal clinically important improvement (defined as the 75th percentile of the change in pain score among patients with a real improvement).

Wording and Response Modalities of the External Anchor	Definition of a Real Improvement	MCII	95% CI
Response to therapy* using Likert 5-point scale	Fair: Reasonable effect, but could be better	–11	–12 to –10
Response to therapy* using Likert 15-point scale	Moderately better	–13	–15 to –11
	Good deal better	–22	–23 to –20
Improvement** (yes or no)	Yes	–18	–19 to –16

* Response to therapy: "How would you rate your response to the NSAID medication you have received for your arthritis for 4 weeks?" ** Improvement: "In your opinion, has the treatment received during the last 4 weeks notably improved your condition?"

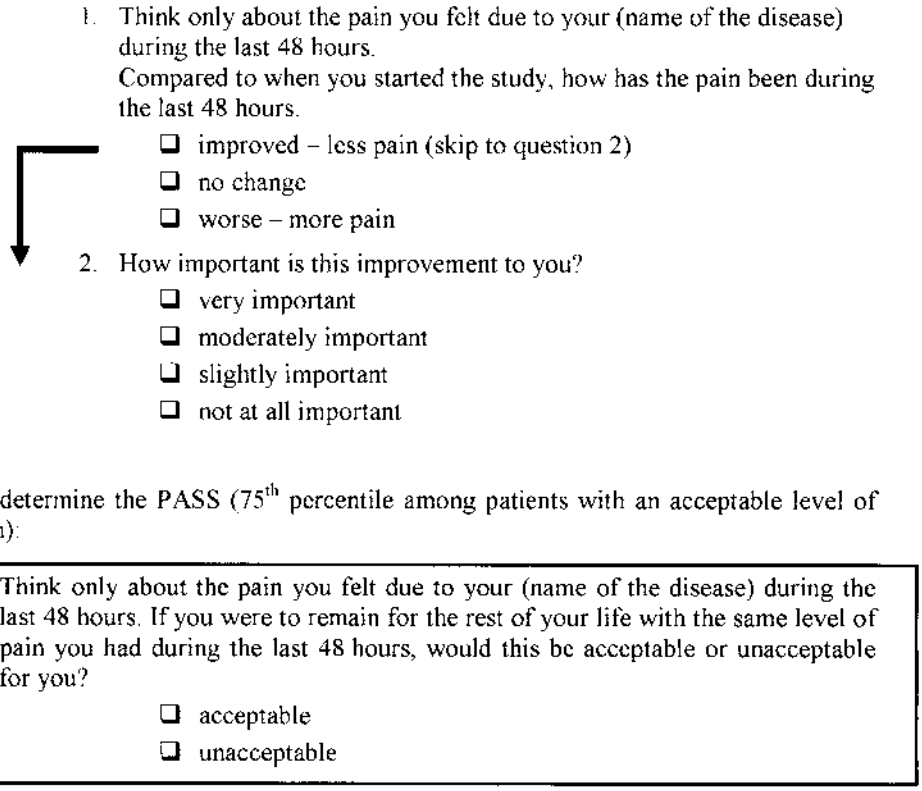

Figure 1. Questions to determine the minimum clinically important improvement (75th percentile among patients moderately improved).

Juniper[12] and Jaeschke[13] are recommended as an external anchor[14]. This external anchor constitutes the first 2 components. The third component is the choice of statistical approach, which can affect the results (Table 2).

Two broad statistical approaches can be considered. The first determines the threshold that best discriminates from the whole sample patients who have improved substantially. Receiver operational characteristic (ROC) curves, classification and regression tree (CART) analysis, or logistic regression can be used. Applying ROC curves allows for choosing the threshold that is the best compromise between sensitivity and specificity (Youden index) for each outcome criterion. But whether sensitivity, specificity, or their sum should be favored depends on the context. Thus, use of the Youden index is arbitrary. This approach has been used by Riddle and associates[15] to determine the MCID for the Roland Morris Back Pain Questionnaire and by Stratford in the Neck Disability Index[16]. The CART is a nonparametric

Table 2. Influence of statistical approach in determining the MCII in pain score in patients with knee osteoarthritis.

Response to Therapy	Cutoff Between:	ROC	CART	75th Percentile	Definition of a Real Improvement
Likert 5-point scale	Poor and fair	–10	–10	–11	Fair
	Fair and good	–28	–34	–20	Good
Likert 15-point scale	Somewhat better and moderately better	–12	–10	–13	Moderately better
	Moderately better and good deal better	–26	–13	–22	Good deal better
Likert 2-point scale	No and yes	–17	–10	–18	Yes

approach for converting continuous variables to categorical variables. This tree-building technique is a form of binary recursive partitioning based on maximal purity.

The above techniques are advantageous because they determine the most accurate change score for patients who have improved. However, because they rely on a binary criterion (improvement yes or no), they do not take into account all the information given by the different response modalities (a global rating on a Likert 5, 7, or 15-point scale).

The second approach is to determine the MCII in the subgroup of patients who experienced an important improvement. In the hip and knee study[9], the MCII was defined as the 75th percentile of change in score among patients whose evaluation of response to therapy on a 5-point Likert scale was "good," because the improvement had to be clinically important. Patients whose evaluation of response to therapy was "excellent" were not included because the target was the minimal change important in the patient's perspective. This definition reflects the target population (75% of the patients that had a good response to therapy had a decreased MCII score).

Methods to determine thresholds should be standardized, and recommendations could be endorsed by an OMERACT module. Studies comparing the accuracy of these different methods could be useful to determine which part of the MCID depends on the statistical approach used.

What Is the Effect of Covariates on the MCII and LDAS/PASS?

The MCII has been demonstrated to vary across tertiles of baseline scores in the subjects of previous studies: in low back pain with use of the Roland Morris Back Pain Questionnaire[15], in chronic obstructive pulmonary disease with the Chronic Respiratory Disease Questionnaire[17], and in hip and knee OA with pain and patient global assessment via the VAS and the WOMAC function subscale[9]. The hip and knee study investigated the effect of several covariates — including age, sex, OA location (hip or knee), and disease duration — on patients' responses. The MCII estimates were not consistently modified. Investigating the potential modifying effects of socioeconomic status and mood (depression and anxiety) in such patient-reported outcomes could be interesting. In another longitudinal study of ankylosing spondylitis, the MCII estimate did not vary for the duration of the study (unpublished data). The only study investigating the influence of potential confounding factors on the PASS[3] found less marked variation in MCII estimates across tertiles of baseline scores, and no effect of age, sex, OA location (hip or knee), or disease duration.

We can conclude that the main confounding factor in the MCII and PASS appears to be the baseline severity of symptoms. Standardization may reduce baseline value effect (to use the relative change instead of the absolute change), but the confounding effect of baseline severity remains (see Table 3). Patients dealing with the most severe symptoms seem to need a greater change to consider themselves improved. This baseline score-related variation may preclude the use of a crude MCII. The patient's initial or previous score should be taken into account when making decisions about important change. Investigators could use the 3 estimates of MCII determined in the tertiles of baseline score to express the changes in terms of important improvement or propose an algorithm based on the baseline value. These solutions meet the recommendation of Crosby and associates[14] for estimating MCID in health-related quality of life criteria, i.e., to anchor the baseline severity of disease in individual patients.

Conclusion

Presenting the results at the level of the individual (proportion of improved patients or patients in an acceptable state) is relevant and provides additional information about effect size. Determining thresholds such as the MCID or LDAS/PASS is important. Because methodological issues affect results (wording, statistical approach, etc.), investigators must standardize methods used to determine thresholds. Threshold values for the MCII and LDAS or PASS for symptomatic outcome criteria in rheumatoid disorders should be adapted from results of studies involving anchor-based methods and experts' opinions. They should also be determined in different datasets involving different clinical environments, languages, and countries.

Table 3. Minimal clinically important improvement (95% CI) in patients with knee osteoarthritis, stratified on the score of interest at baseline divided into tertiles (defined as the 75th percentile of the change in score among patients whose evaluation of response to therapy was "good," in terms of 3 patient-related outcomes: pain, as assessed on visual analog scale (VAS), global assessment of disease status on VAS, or the Western Ontario and McMaster Universities (WOMAC) function subscale.

Measure	Absolute Change Baseline Score Tertile			Relative Change Baseline Score Tertile		
	Low	Intermediate	High	Low	Intermediate	High
Pain, VAS, 0–100 mm	–11 (–13 to –9)	–27 (–30 to –25)	–37 (–38 to –35)	–29 (–34 to –24)	–47 (–53 to –41)	–51 (–56 to –46)
Patient's global assessment, VAS, 0–100 mm	–6 (–9 to –4)	–25 (–27 to –22)	–43 (–47 to –39)	–20 (–26 to –13)	–43 (–48 to –32)	–58 (–64 to –50)
WOMAC function subscale, 0–100	–5 (–7 to –4)	–12 (–13 to –10)	–20 (–23 to –18)	–22 (–27 to –17)	–26 (–29 to –24)	–33 (–36 to –30)

REFERENCES

1. Wells G, Beaton D, Shea B, et al. Minimal clinically important differences: review of methods. J Rheumatol 2001;28:406-12.
2. Wells G, Boers M, Shea B, et al. MCID/Low Disease Activity State Workshop: Low disease activity state in rheumatoid arthritis. J Rheumatol 2003;30:1110-1.
3. Tubach F, Ravaud P, Baron G, et al. Evaluation of clinically relevant states in patient-reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. doi: 10.1136/ard.2004.023028. Ann Rheum Dis 2005;64:34-7.
4. Wells G, Anderson J, Beaton D, et al. Minimal clinically important difference module: summary, recommendations, and research agenda. J Rheumatol 2001;28:452-4.
5. Kirwan JR. Minimum clinically important difference: the crock of gold at the end of the rainbow? J Rheumatol 2001;28:439-44.
6. Goldsmith CH, Boers M, Bombardier C, Tugwell P. Criteria for clinically important changes in outcomes: development, scoring and evaluation of rheumatoid arthritis patient and trial profiles. OMERACT Committee. J Rheumatol 1993;20:561-5.
7. Felson DT, Anderson JJ, Boers M, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995;38:727-35.
8. Wells G, Anderson J, Boers M, et al. MCID/Low Disease Activity State Workshop: Summary, recommendations, and research agenda. J Rheumatol 2003;30:1115-8.
9. Tubach F, Ravaud P, Baron G, et al. Evaluation of clinically relevant changes in patient-reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. doi: 10.1136/ard.2004.022905. Ann Rheum Dis 2005;64:29-33.
10. Wells G, Boers M, Shea B, et al. Minimal disease activity for rheumatoid arthritis: A preliminary definition. J Rheumatol 2005;32:2016-24.
11. Falgarone G, Zerkak D, Baver C, Messow M, Dougados M. How to define a minimal clinically individual state with pain VAS in daily practice for patients suffering from musculoskeletal disorders. Clin Exp Rheumatol 2005;23:235-8.
12. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific quality of life questionnaire. J Clin Epidemiol 1994;47:81-7.
13. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials 1989;10:407-15.
14. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol 2003;56:395-407.
15. Riddle DL, Stratford PW, Binkley JM. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: Part 2. Phys Ther 1998;78:1197-207.
16. Stratford PW, Riddle DL, Binkley JM, Spadoni G, Westaway MD, Padfield B. Using the neck disability index to make decisions concerning individual patients. Physiother Can 1999;Spring:107-12.
17. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. J Clin Epidemiol 1999;52:861-73.