Editorial

MCID — The Minimal Clinically Important Difference Assigns Significance to Outcome Effects

At medical school we were taught that 70–80% of the information that leads to diagnosis comes from the patient's anamnesis. The severity of the subjective symptoms and disabilities drives the patient to seek medical help and has a major influence on treatment and intervention, despite "objective" findings such as alterations on a radiograph, for example, in a knee affected by osteoarthritis (OA)[1]. Based on this insight, a huge number of patient-rated outcome instruments have been developed in the past 2 decades[2]. Nevertheless, the significance of therapeutic effects is still quantified by statistical methods alone in many study reports, especially in pharmacological ones, even if the effects are labeled as "clinically significant"[3].

Beyond statistically and distribution-based quantification of effect significance, the dimension of an effect's importance and significance, which includes the patient's subjective perception of pain and function, reaches a higher sphere because it is closer to the central subject of interest in medicine, the patient. "It is recommended that the patient's perspective be given the most weight, because these are patient-related outcome measures, although the clinician's perspective is considered important as well"[4]. The founders of the concepts developed consequently to give this alternative meaning for outcome effects were Jaeschke and Redelmeier[5,6].

Jaeschke was the first investigator to ask patients to rate their subjectively perceived change of health or symptoms between baseline and followup[5]. The responses on this "transition" item were related to the score differences of an outcome instrument within the same time period — the so-called anchor-based technique. The mean score difference (between baseline and followup) of the transition group responses "almost the same," "a little better," and "somewhat better" was defined as the minimal clinically important difference (MCID) for improvement — I call it the "crude" mean change method[5].

In rheumatology, Redelmeier developed this concept further to become the so-called "corrected" mean change method, the one most frequently applied in investigative studies[6]. The mean score difference (between baseline and followup) of the

"somewhat better" group ("a little" and "somewhat" were collapsed into "somewhat") was reduced by that of the "almost the same" group "to adjust for possible bias in ratings…"[6].

This estimate for an effect is "the smallest difference in score in the domain of interest that patients perceive as beneficial and that would mandate, in the absence of troublesome side effects and excessive costs, a change in the patient's management"[5]. The relationship of the anchor-based MCID to distribution-based smallest (statistically) significant differences has been illustrated in several studies and reviews[2,4,7,8,9]. The potential for conflicts when one concept competes against another is inherent and has been highlighted by many authors[2,4,7,8,9].

An early study of outcomes for hip and knee OA patients after inpatient rehabilitation established an MCID for improvement of 8.3 score points (scale 0–100) on the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain scale, and 8.0 on the WOMAC function scale[10]. However, there are few investigative studies in the literature that have evaluated the size of MCID in OA settings. The fundamental handbook of outcome instruments in rheumatology stated in 2011 that the MCID of the Knee injury and Osteoarthritis Outcome Score (KOOS) had not yet been determined[2].

Thus, the work of Mills, *et al* in this issue of *The Journal* is all the more deserving of plaudits[11]. Their study examined the MCID, or as the authors prefer to call it, the MID (minimal important difference) for improvement and worsening on the KOOS, which incorporates the items of the WOMAC. The 280 included patients with knee OA scored at least 30/100 mm on pain visual analog scale and underwent individually tailored conservative (nonsurgical) outpatient management.

Specific positive characteristics of the study are that the M(C)ID were stratified by followup (12, 26, and 52 weeks) and adjusted by baseline variables sex, age, pain severity, 6-min walking test, waitlisted for joint replacement (yes/no), and unilateral/bilateral symptoms. The effects of 2 different anchors, walking level and knee symptoms, were examined.

The KOOS score differences were correlated to the transition/anchor responses, and M(C)ID analysis was performed only if the correlation was ≥ 0.40[4]. The M(C)ID were determined by the 2 mean change methods as well as 2 receiver-operation characteristics curve (ROC) analyses. The first ROC, the "Youden" method, maximizes the sum of sensitivity plus specificity, and the second, the 80% specificity method[5,6,9], works with a fixed minimum of specificity.

Analysis partly revealed wide variation of M(C)ID depending on the 4 chosen methods. By definition, the crude and the corrected mean change methods differed by the mean score difference of the "not changed" response group. In some analyses, it was close to zero [e.g., M(C)ID for improvement of function]; in others, not [e.g., M(C)ID for improvement of pain at 26 weeks]. The 2 ROC methods resulted in similar levels in some analyses and in quite different ones in others. The differing results highlight the effects and importance of the *a priori* chosen method to determine M(C)ID and the underlying assumptions. This issue is discussed at great length as part of the study of Mills, *et al*[11].

The 2 mean change methods provide a measurement that is obtained from a sample of several individuals and is a mean value. In contrast, the M(C)ID obtained by the ROC methods is a single empirical result. From all the individual score differences of the sample lying on the ROC curve, the ROC method chooses one for which the sum of the sensitivity plus the specificity to differentiate "slightly better" from "not changed" is maximal[9]. This characteristic of a single point estimate that maximizes differentiation questions by the ROC method is that it provides a valid, minimal effect estimate of perceived changes in health.

Beside the chosen concepts, additional methods to quantify, for example, the use of relative instead of absolute effect parameters may yield M(C)ID that are easier to apply to future studies. The use of percentages of the baseline score is 1 possibility, as also discussed in the study by Mills, *et al*[11]. Another might be the use of effect sizes, i.e., the mean change (between baseline and followup) in either absolute score points divided by the SD of the whole group at baseline or of the group's score differences or the differences pooled between 2 groups[4,12].

The study by Mills, *et al* demonstrates that many different attitudes, assumptions, and theoretical considerations exist, which results in different concepts and empirical quantifications of M(C)ID[11]. Confusion is complete if anchor-based methods are mixed up with distribution-based ones that use statistical computation methods alone[4,9]. Discussion of which method is best for which condition is ongoing in the scientific world[4].

One solution could be based on a consensus to return to the patients, who are the focus of our work, and consequently, to the original concepts of Jaeschke and Redelmeier[5,6]. Subtraction of the mean score difference of the "not changed" group from that of the "slightly better/worse" group according to Redelmeier might be elaborated further and be less biased than the crude estimate of Jaeschke, who, inciden-tally, "did not claim to actually derive a truly 'minimal' difference estimate"[4]. Multivariate adjustment for baseline variables and characterizing/determination/quantification M(C)ID by sample-based effect sizes offer possibilities that might reduce different sorts of bias to obtain valid M(C)ID rather than single point estimates.

FELIX ANGST, MD, MPH,
Head of Research Department,
RehaClinic,
Bad Zurzach, Switzerland
Address correspondence to Dr. F. Angst, RehaClinic, Quellenstr. 34, 5330 Bad Zurzach, Switzerland; E-mail: fangst@vtxmail.ch, f.angst@rehaclinic.ch

REFERENCES
1. Angst F, Stucki G, Aeschlimann A. Quality of life assessment in osteoarthritis. Expert Rev Pharmacoeconomics Outcomes Res 2003;3:623-36.
2. Katz PP, editor. Patient outcomes in rheumatology, 2011. A review of measures. Arthritis Care Res 2011;63 Suppl:S1-S490.
3. Bannuru RR, Schmid CH, Kent DM, Vaysbrot EE, Wong JB, McAllindon TE. Comparative effectiveness of pharmacologic interventions for knee osteoarthritis. A systematic review and network analysis. Ann Intern Med 2015;162:46-54.
4. Revicki D, Hays RD, Cella DE, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61:102-9.
5. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Controlled Clin Trials 1989;10:407-15.
6. Redelmeier DA, Lorig K. Assessing the clinical importance of symptomatic improvements. An illustration in rheumatology. Arch Intern Med 1993;153:1337-42.
7. Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. Arthritis Rheum 2001;45:384-91.
8. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J 2007;7:541-6.
9. De Vet HC, Terluin B, Knol DL, Roorda LD, Mokkink LB, Ostelo RW, et al. Three ways to quantify uncertainty in individually applied "minimally important change" values. J Clin Epidemiol 2010; 63:37-45.
10. Angst F, Aeschlimann A, Michel BA, Stucki G. Minimal clinically important rehabilitation effects in patients with osteoarthritis of the lower extremities. J Rheumatol 2002;29:131-8.
11. Mills K, Naylor J, Eyles J, Roos E, Hunter D. Examining the minimal important difference of patient reported outcome measures for individuals with knee osteoarthritis: a model using the knee injury and osteoarthritis outcome score. J Rheumatol 2016;43:395-404.
12. Borenstein M. Effect sizes for continuous data. In: Cooper H, Hedges LV, Valentine JC, editors. The handbook of research synthesis and meta-analysis. 2nd ed. New York: Russell Sage Foundation; 2009:222-36.