PADUA DAYS ON MUSCLE AND MOBILITY MEDICINE 2026

International Conference

March 3rd to 6th Euganean Thermae and Padua, Italy

PADUA DAYS ON MUSCLE AND MOBILITY MEDICINE 2026 ABSTRACT N. 083

LUDWIG BOLTZMANN INSTITUTE WORKSHOP ON REHABILITATION ASSESSMENTS

MEASURING WHAT MATTERS: THE SIGNIFICANCE OF BASELINE, CHANGE, AND OUTCOME IN REHABILITATION RESEARCH

Ferdinand Prüfer[1|2], Spela Matko[1|3], Stefan Löfler[1|4], Michael J. Fischer[1|5], Vincent Grote[1|2]

Ludwig Boltzmann Institute for Rehabilitation Research, Vienna, Austria;Medical University of Graz, Graz, Austria; tnstitute for Outcomes Research, Medical University of Vienna, Vienna, Austria; [4]Physiko- und Rheumatherapie, Institute for Physical Medicine and Rehabilitation, St Poelten, Austria; [5]Rehabilitation Center Kitzbühel, Kitzbühel, Austria.

Rehabilitation medicine relies on patient-reported and clinician-reported outcome measures (PROMs and CROMs) to assess patient progress, guide clinical decision-making, and inform policy. Because rehabilitation primarily evaluates trajectories, change scores are central. However, interpreting these changes can be challenging, as not every detected change is relevant to clinical practice. To better understand individual patients' progression over the course of rehabilitation, clear reference points are essential. On the one hand, we need normative data, which describe what is usual in a defined population and context – i.e., the typical level for comparable patients (for example, functional status, age, sex). On the other hand, we need instruments that identify meaningful changes. Normative data summarize what is typical in a defined population at a given time and provide reference values to identify deviations and set care benchmarks (4). They may be national, regional, local, or disease-/setting-specific, but are interpretable only for individuals from the same reference population. Comparing PROMs or CROMs with norms helps quantify disease burden and inform planning. However, broad national norms often lack disease specificity and may not reflect local patient characteristics or rehabilitation protocols. In rehabilitation, change is typically expressed as absolute (post–pre) or relative (% from baseline) differences. These derived endpoints can lead to different statistical conclusions from the same data and may differ in power, so the choice should match the analytic purpose. Absolute change is reported in the instrument's native units, is straightforward to interpret, and is symmetric with respect to direction (the difference from A to B equals the negative of B to A). Relative change is unit-free and facilitates comparisons across measures or settings, but it has well-known limitations: it is asymmetric (A being 5% higher than B does not imply B is 5% lower than A), shows strong baseline dependence (identical absolute improvements produce larger percentages at lower baselines and smaller percentages at higher baselines), and becomes unstable near low or floor values, which can skew distributions and reduce statistical power in heterogeneous samples (3). Given these properties, absolute change is often preferable as the primary summary of trajectories in rehabilitation. Nevertheless, statistical measurable change (e.g. significance) does not equal meaningful change, which requires a clinical interpretive threshold.The minimal clinically important difference (MCID) is the smallest change in an outcome that patients perceive as beneficial and that could prompt a change in care (2). As a patient-derived threshold, the MCID complements statistical significance by indicating when change is clinically relevant. It also supports practical decisions – defining responders, setting goals (and expectations) with patients, and communicating benefits across instruments and settings. However, there is a lack of well-derived and properly reported MCIDs for many commonly used outcome measures. Moreover, MCIDs are population- and method-specific: different estimation approaches can yield different thresholds, and literature-derived values are often misapplied when transferred across populations or used as fixed individual cut-offs. MCIDs should therefore be interpreted within context and alongside measurement properties of the instrument. The size and meaning of change—and the likelihood of exceeding an MCID—depend strongly on the baseline level of the patient. For a given absolute improvement, percentage gains look larger at lower starting values and smaller at higher ones; more importantly, patients beginning with greater impairment often require larger absolute improvements to experience a change they judge meaningful. A clinical illustration: lowering systolic blood pressure from 190 to 170 mmHg may be perceived as more consequential than a reduction from 130 to 110 mmHg, even though both reflect 20 mmHg of absolute change. While sociodemographic and clinical factors also play a role, baseline status typically exerts the largest influence (1). Consequently, "one-size-fits-all" thresholds are inadequate; applying norms or MCIDs to individual patients should account for disease context, baseline level, and relevant patient characteristics.To make MCIDs clinically applicable at the individual level, thresholds should be conditional on baseline and, where appropriate, other covariates (e.g. age, sex). Two pragmatic implementations are: (a) baseline-stratified tables and (b) model-based formulas that map an individual's baseline value to a personalized threshold. Although some methodological building blocks for bias-aware, covariate-adjusted thresholds exist (Terluin et al., 2017), their development and application remain sparse and confined to a few instruments. Most outcome measures still

lack appropriately derived and adjusted MCIDs. The Patient Acceptable Symptom State (PASS) defines a level of symptoms or function that patients consider satisfactory at a given time point – i.e., "how good is good enough?" – rather than the amount of change required (MCID). As a state-based endpoint, PASS is intuitive for shared decision-making, directly interpretable at a single time point, and useful when baseline varies widely or when small changes near a desirable state matter (6). It can also complement responder analyses by indicating whether patients have reached a target condition, regardless of how far they traveled to get there. Compared with MCID, PASS (i) is less sensitive to baseline-driven artifacts in percentage change, (ii) aligns well with clinical goals framed as acceptable functioning or pain levels, and (iii) supports patient-centered counseling about readiness for discharge or return to activity. However, PASS (i) remains population- and context-specific and requires carefully validated anchors, (ii) may overlook meaningful improvement if the acceptable (possible) state is not reached (useful change but "non-PASS"), and (iii) can be affected by response shift or adaptation over time. MCID, by contrast, (i) quantifies meaningful change – valuable for trials and quality improvement – but (ii) is method-dependent and often misapplied as a fixed individual cut-off. In practice, clinicians should use them together: MCID to judge whether improvement is meaningful, and PASS to judge whether the current state is acceptable. This framework organizes rehabilitation evaluation around three pillars: PROMs and CROMs. locate where patients stand (normative reference data), and decide what counts (dynamic MCIDs & PASS). This approach replaces universal cut-offs with interpretable, context-specific targets, mitigates baseline- and method-driven distortions, and strengthens the clinical meaning of change. In practice, it supports clearer goal setting, more patient-centered communication, and fairer benchmarking across instruments and settings.

Keywords: *rehabilitation outcomes, change scores, normative reference data, MCID, PASS*