The Journal of Rheumatology 2022;49:615–21
doi:10.3899/jrheum.210543
First Release May 1 2022

The Journal of Rheumatology

WOMAC Meaningful Within-patient Change: Results From 3 Studies of Tanezumab in Patients With Moderate-to-severe Osteoarthritis of the Hip or Knee

Philip G. Conaghan[1], Robert H. Dworkin[2], Thomas J. Schnitzer[3], Francis Berenbaum[4], Andrew G. Bushmakin[5], Joseph C. Cappelleri[5], Lars Viktrup[6], and Lucy Abraham[7]

ABSTRACT. ***Objective.*** To define meaningful within-patient change (MWPC) in the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC).

Methods. Data were analyzed separately from 3 phase III clinical trials (ClinicalTrials.gov: NCT02697773, NCT02709486, NCT02528188) of tanezumab, a novel treatment intended for the relief of signs and symptoms of moderate-to-severe osteoarthritis (OA), administered subcutaneously every 8 weeks. Patients with moderate-to-severe OA of the hip or knee completed the WOMAC and patient global assessment of OA (PGA-OA) at regular timepoints. A repeated measures longitudinal model with change in WOMAC Pain, Physical Function, or Stiffness domain score as the outcome and change in PGA-OA as the anchor was used to establish MWPC for WOMAC domains.

Results. In the 3 studies, there were 688, 844, and 2948 subjects available for analyses, respectively. Analysis showed that a linear relationship between changes in WOMAC domains and changes in PGA-OA was supported and justified. Moreover, the relationships between these changes were very similar for 2 trials and close for the third. The estimated MWPC for the 3 WOMAC domains were from 0.84–1.16 (0–10 numerical rating scale) and from 12.50–16.23%, depending on study and domain, that corresponded to a 1-category change on PGA-OA. For a 2-category change those values were from 1.68–2.31 and from 25.01–32.46%, respectively.

Conclusion. These results establish MWPCs for WOMAC domains, at the individual patient level, for patients with moderate-to-severe OA of the hip or knee. [ClinicalTrials.gov: NCT02697773, NCT02709486, and NCT02528188]

Key Indexing Terms: clinical trials, osteoarthritis, outcomes, pain

Clinical trial outcomes are typically determined by statistically significant mean differences between treatment and control groups; however, these "between-group" analyses do not indicate whether individual patients have experienced clinically meaningful benefits. To aid interpretation of the clinical meaningfulness of responses to treatment, it is important to determine thresholds for outcome measures. Such thresholds make it possible to assess group differences in the percentages of participants who have meaningful improvements after starting treatment.

Over 30 years ago, the minimal clinically important difference (MCID) was defined as "the smallest difference in score in the domain of interest which patients perceive as beneficial and would mandate, in the absence of troublesome side effects and excessive costs, a change in the patient's management".[1] Although this definition of MCID refers to changes that patients perceive as meaningful for themselves, typically established using a global anchor rating of disease status or change, such MCIDs have been used in the interpretation of group differences observed

This study was sponsored by Pfizer Inc and Eli Lilly and Company. PGC is supported in part through the UK National Institute for Health Research (NIHR) Leeds Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the UK NHS, the NIHR, or the UK Department of Health.

[1]P.G. Conaghan, MB BS, PhD, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, and NIHR Leeds Biomedical Research Centre, Leeds, UK; [2]R.H. Dworkin, PhD, University of Rochester, Rochester, New York, USA; [3]T.J. Schnitzer, MD, PhD, Northwestern University, Evanston, Illinois, USA; [4]F. Berenbaum, MD, PhD, Sorbonne Université, Paris, France; [5]A.G. Bushmakin, MS, J.C. Cappelleri, PhD, Pfizer Inc, Groton, Connecticut, USA; [6]L. Viktrup, MD, PhD, Eli Lilly and Co., Indianapolis, Indiana, USA; [7]L. Abraham, MSc, CPsychol, Pfizer Ltd, Surrey, UK.

PGC is a paid consultant for AbbVie, BMS, Eli Lilly, EMD Serono, Flexion Therapeutics, Galapagos, Gilead, GSK, Janssen, Novartis, Pfizer, Regeneron, and Samumed. RHD is a paid consultant for Abide, Acadia, Analgesic Solutions, Asahi Kasei, Biogen, Centrexion, Clexio, Decibel, Eli Lilly, Glenmark, Hope, Lotus, Mainstay, Merck, Neurana, NeuroBo, Novaremed, Novartis, Pfizer, Regenacy, Sanifit, Scilex, Semnur, Sollis, Vertex, and Vizuri. TJS is a paid consultant for AstraZeneca, Eli Lilly, Pfizer, and Regeneron. FB is a paid consultant for Eli Lilly and Pfizer. AGB, JCC, and LA are employees of and stockholders in Pfizer. LV is an employee of and stockholder in Eli Lilly.

Address correspondence to L. Abraham, Pfizer Ltd, 1-G Walton Oaks, Dorking Road, Walton on the Hill, Tadworth, Surrey KT20 7NS, UK. Email: Lucy.Abraham@pfizer.com.

Accepted for publication February 9, 2022.

© 2022 The Journal of Rheumatology. This is an Open Access article, which permits use, distribution, and reproduction, without modification, provided the original article is correctly cited and is not used for commercial purposes.

Downloaded on October 10, 2026 from www.jrheum.org

in clinical trials.[2,3] However, there are typically different considerations involved when interpreting what patients consider important improvements vs what should be considered an important group difference in a clinical trial.

The distinction between these 2 different concepts of clinical importance has been recognized for many years across a variety of therapeutic areas. It has typically been concluded that the clinical importance of group differences should not be based solely on the magnitude of the within-patient improvements that patients (or clinicians) consider important but should rather be based on the broader context of the disease and its available treatments, along with a careful evaluation of risk vs benefit by relevant stakeholders.[4,5] Indeed, in a landmark article, Guyatt and colleagues emphasized, "Clinicians and investigators tend to assume that if the mean difference between a treatment and a control is appreciably less than the smallest change that is important, then the treatment has a trivial effect. This may not be so. Let us assume that a randomised clinical trial shows a mean difference of 0.25 in a questionnaire in which the minimal important difference is 0.5. It might be concluded that the difference is unimportant and that the result does not support giving the treatment. This interpretation assumes that every patient treated scored 0.25 better than they would have done had they received the control and ignores the possibility that treatment might have a heterogeneous effect."[6]

Because of these considerations, use of the term *MCID* can be confusing when the definition is not very clearly specified and the difference between individual improvements and group differences is not clarified. Given this potential for misunderstanding, a recent US Food and Drug Administration (FDA) guidance on patient-focused drug development emphasized the difference between group- and individual-level changes, and highlighted the importance of defining what constitutes meaningful within-patient change (MWPC) in regulatory submissions.[7] Just as thresholds for meaningful individual improvements should not necessarily be used to evaluate group differences, the FDA emphasized that "a treatment effect is different than a meaningful within-patient change. The terms minimally clinically important difference (MCID) and minimum important difference (MID) do not define meaningful within-patient change if derived from group level data and therefore should be avoided."[7]

The FDA guidance recommended the use of methods that utilize an external anchor measure to define the criteria for MWPC of an outcome measure at the individual level.[7] Anchor measures should be plainly understood and easier to interpret than the outcome measure. Recommended anchors include current-state global impression of severity scales. These scales may be preferable, as they avoid the recall bias associated with global impression of change scales and can also be used to assess change from baseline.[7,8] Given a large range of thresholds reported in the literature, it is also important to evaluate the responder definition of an outcome for the target population of an intervention.

The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) is a disease-specific measure of osteoarthritis (OA) symptoms (pain and stiffness) and functional impairment that is regularly used in clinical trials. In order to facilitate the interpretation of OA clinical trial endpoints, it is important to determine thresholds for WOMAC that would indicate MWPC.

Previous attempts to define MWPC through metaanalysis of clinical trials in heterogeneous chronic pain conditions identified thresholds of 10–20% reductions in pain as minimally important, ≥ 30% reductions as moderately important, and ≥ 50% reductions as substantial.[4,9] A systematic literature review examined anchor-based clinically meaningful changes in WOMAC scores reported for patients with OA.[10] Values for clinically important differences in WOMAC scores were between 0.3–1.3 for patients with OA of the hip and knee, between 0.5–3.4 for those with OA of the knee, and between 0.3–3.6 for those with OA of the hip, depending on the domain assessed (original standardized 0–100 scale converted to 0–10 scale).

The patient global assessment of OA (PGA-OA) is a current-state global impression of disease severity scale regularly used in OA clinical trials in combination with WOMAC.[11,12] Given the large range of values reported for clinically meaningful changes in WOMAC, our aim was to estimate MWPC in WOMAC by examining the relationship between change in the outcomes of WOMAC domains and change in the anchor of PGA-OA using data from 3 trials that included patients with moderate-to-severe OA of the hip or knee.

METHODS

Statement of ethics and consent. The protocol for each clinical trial (ClinicalTrials.gov: NCT02697773, NCT02709486, and NCT02528188) was approved by an institutional review board or independent ethics committee at each participating investigational center. The studies were conducted in compliance with the ethical principles of the Declaration of Helsinki and Good Clinical Practice Guidelines. All patients provided written informed consent before entering the studies.

Patients. Data were analyzed separately from 3 phase III clinical trials of tanezumab, a novel treatment intended for the management of OA pain in adult patients for whom the use of other analgesics is ineffective or not appropriate. Study 1 (ClinicalTrials.gov: NCT02697773) was a 16-week randomized, double-blind, placebo-controlled trial.[13] Patients in study 1 received subcutaneous (SC) placebo or tanezumab 2.5 mg at baseline and week 8, or tanezumab 2.5 mg at baseline and 5 mg at week 8. Study 2 (NCT02709486) was a 24-week randomized, double-blind, placebo-controlled trial.[14] Patients in study 2 received SC placebo or tanezumab 2.5 mg or 5 mg at baseline and weeks 8 and 16. Study 3 (NCT02528188) was a 56-week randomized, double-blind, double-dummy, active-controlled trial.[15] Patients in study 3 received oral nonsteroidal antiinflammatory drugs twice daily or SC tanezumab 2.5 mg or 5 mg every 8 weeks, with oral and SC study medication controls.

For all studies, patients aged ≥ 18 years with a BMI of ≤ 39 kg/m^2 and a diagnosis of OA (Kellgren-Lawrence [KL] grade ≥ 2) of the hip or knee were eligible to enroll. Patients were required to have WOMAC Pain scores ≥ 5 in the index joint at baseline, defined as the most painful joint at baseline with a qualifying WOMAC Pain score and KL grade, as confirmed by the central reader's assessment. In all studies, patients were also required to have a WOMAC Physical Function score ≥ 5 in the index joint and a PGA-OA of fair, poor, or very poor at baseline.

Measures. WOMAC Index Version 3.1 numeric rating scale (NRS) consisted of 24 items assessed using a 0–10 NRS (0 = no pain/stiffness/difficulty to 10 = extreme pain/stiffness/difficulty), with higher scores indicating worse outcomes.[16] The WOMAC Pain (5 items), Stiffness (2 items),

Downloaded on October 10, 2026 from www.jrheum.org

and Physical Function (17 items) subscales measured the pain, stiffness, or difficulty experienced in performing common tasks during a 48-hour recall period. Scores for total WOMAC and each subdomain were calculated from the average score of the component questions.

PGA-OA was a single question: "Considering all the ways your OA in your hip/knee affects you, how are you doing today?" PGA-OA was measured on a 5-point Likert scale, with higher scores indicating worse symptoms (1 = very good [asymptomatic and no limitation of normal activities] to 5 = very poor [very severe symptoms that are intolerable and inability to carry out all normal activities]; Supplementary Table 1, available with the online version of this article). Patients were only asked to complete WOMAC and PGA-OA for their index hip or index knee, depending on the initial assessment for each subject.

WOMAC and PGA-OA were completed electronically on tablets at site visits. In study 1, all WOMAC subscales were completed at screening, baseline, and weeks 2, 4, 8, 12, 16, and 24; PGA-OA was completed at baseline and weeks 2, 4, 8, 12, 16, and 24. In study 2, WOMAC Pain was completed at screening; all WOMAC subscales and PGA-OA were completed at baseline and weeks 2, 4, 8, 12, 16, 24, and 32. In study 3, all WOMAC subscales were completed at screening, baseline, and weeks 2, 4, 8, 16, 24, 32, 40, 48, 56, and 64; PGA-OA was completed at baseline and weeks 2, 4, 8, 16, 24, 32, 40, 48, 56, and 64.

Statistical analyses. Changes from baseline in WOMAC and PGA-OA were calculated as the postbaseline value minus the baseline, with lower change score being more favorable. Negative changes therefore represented improvement in both measures.

A repeated measures longitudinal model with change in WOMAC domain score as the outcome and change in PGA-OA as the anchor measure was used to establish MWPC for WOMAC domains.[17,18,19,20,21] In such a model, change from baseline in a WOMAC domain score was taken as a dependent variable using all available data from weeks 2–24 in study 1, from weeks 2–32 in study 2, and from weeks 2–64 in study 3. The model contained data combined across treatment groups.

To study the appropriateness of the linear approximation of the relationship between predictor and outcome, we also studied the model when a predictor was used as a categorical variable. This version of the model does not impose any functional relationship, linear or otherwise, between an outcome and a predictor.

MWPC was evaluated using a 1-category change and, separately, a 2-category change on the anchor measure. The repeated measures longitudinal model calibrated the relationship between the outcome variable and anchor predictor by taking the difference in mean change in outcome scores between adjacent categories of the change in the anchor predictor.[22] The theoretical range for original unit changes in WOMAC scores is from −10 to 10. Conversion to percentage change leads to a range from −100% to, theoretically, plus infinity for changes in WOMAC domains scores.

In addition to producing estimates for MWPC for all patients within each study, subgroups of patients based on index joint were separately analyzed to investigate any potential differences in MWPC for patients with an index hip or knee.

The anchor-based repeated measures longitudinal model was supplemented with empirical cumulative distribution function (eCDF) analyses, in accordance with FDA recommendations.[7] eCDFs were produced at the primary timepoints of week 16 in studies 1 and 3 and of week 24 in study 2.

RESULTS

The relationship between changes in PGA-OA and changes in WOMAC domains was analyzed in data collected from a total of 4480 patients, corresponding to 688, 844, and 2948 patients in studies 1, 2, and 3, respectively. Patient demographics and disease status were similar across the 3 studies (Table 1).

Using change in PGA-OA as a continuous or categorical variable revealed a robust relationship with change in all WOMAC domains across all 3 studies (Figure 1; Supplementary Figures 1–2, available with the online version of this article). The correlations between changes in WOMAC Pain and Physical Function and changes in PGA-OA were approximately 0.6 for studies 1 and 3, and approximately 0.5 in study 2 (Table 2). Correlations between changes in WOMAC Stiffness and changes in PGA-OA were approximately 0.5 in studies 1 and 3, and approximately 0.4 in study 2. These data showed that a linear relationship between changes in WOMAC domains and changes in PGA-OA was supported and justified.

The relationship between changes in WOMAC domains and PGA-OA was examined across the 3 studies (Figure 2; Supplementary Tables 2A–C, available with the online version of this article). The estimated MWPC for the 3 WOMAC domains were 0.84–1.16 (original units) and 12.50–16.23%, depending on study and domain, that corresponded to a 1-category change on PGA-OA (Tables 3A,B). For a 2-category change the corresponding values were 1.68–2.31 (original units) and 25.01–32.46% (Tables 3A,B). The estimated MWPC calculated separately for patients with an index hip or knee were not appreciably different from the pooled analyses or between subgroups of patients with an index hip or knee (Supplementary Tables 3–4).

To supplement the anchor-based method to estimate MWPC, eCDF curves for changes from baseline in WOMAC domains categorized by change in PGA-OA were generated. eCDFs for changes in WOMAC Pain in studies 1 and 3 showed a clear separation of curves for categories of PGA-OA changes with a sufficient number of available observations per

Table 1. Baseline demographics and disease status of patients in each study.[a]

	Study 1, n = 696	Study 2, n = 849	Study 3, n = 2996
Age, yrs	60.8 (9.6)	64.9 (9.4)	60.6 (9.4)
Female, n, (%)	453 (65.1)	587 (69.1)	1953 (65.2)
Index joint, n, (%)			
Hip	102 (14.7)	114 (17.0)	443 (14.8)
Knee	594 (85.3)	705 (83.0)	2553 (85.2)
Kellgren-Lawrence grade of index joint, n, (%)			
0	0	2 (0.2)	5 (0.2)
1	1 (0.1)	0	7 (0.2)
2	184 (26.5)	166 (19.6)	892 (29.8)
3	304 (43.7)	375 (44.2)	1425 (47.6)
4	206 (29.6)	306 (36.0)	667 (22.3)
Disease duration, yrs	9.3 (8.4)	7.5 (7.0)	8.8 (8.3)
WOMAC Pain	7.2 (1.2)	6.6 (0.9)	7.0 (1.1)
WOMAC Physical Function	7.3 (1.1)	6.7 (0.9)	7.1 (1.1)
PGA-OA	3.5 (0.6)	3.6 (0.6)	3.5 (0.6)

Values are expressed as mean (SD) unless otherwise indicated. [a] Data were analysed from 688, 844 and 2948 patients from studies 1, 2, and 3, respectively. Patients without data available at all timepoints in the respective studies were excluded from the analyses. PGA-OA: patient global assessment of osteoarthritis; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

Downloaded on October 10, 2026 from www.jrheum.org

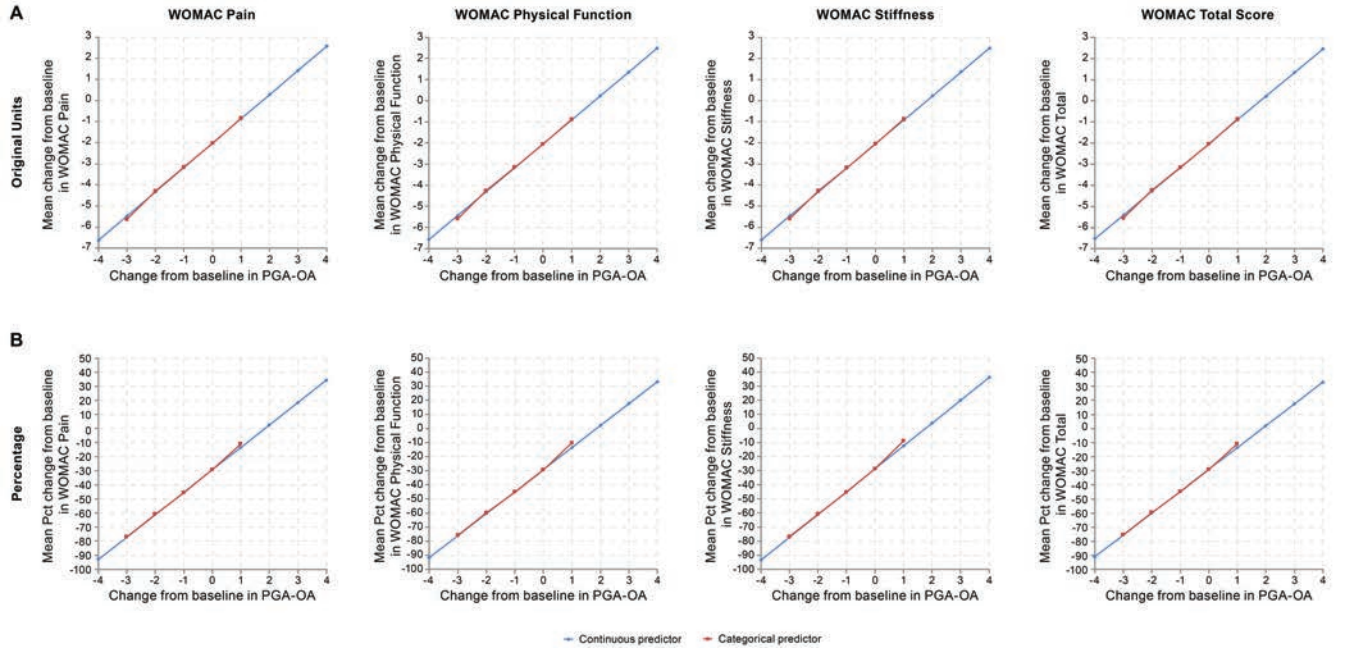

Figure 1. Relationship between change in WOMAC domains and PGA-OA in study 3. Using PGA-OA as a continuous variable (blue lines), which implies a linear relationship with the outcome, or as a categorical variable (red lines), which does not impose any functional relationship, revealed an approximately linear relationship with all WOMAC domains in study 3. This relationship is observed when change in WOMAC is reported (A) in original units or (B) as a percentage. A negative change represents an improvement in both measures. Pct: percent; PGA-OA: patient global assessment of osteoarthritis; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

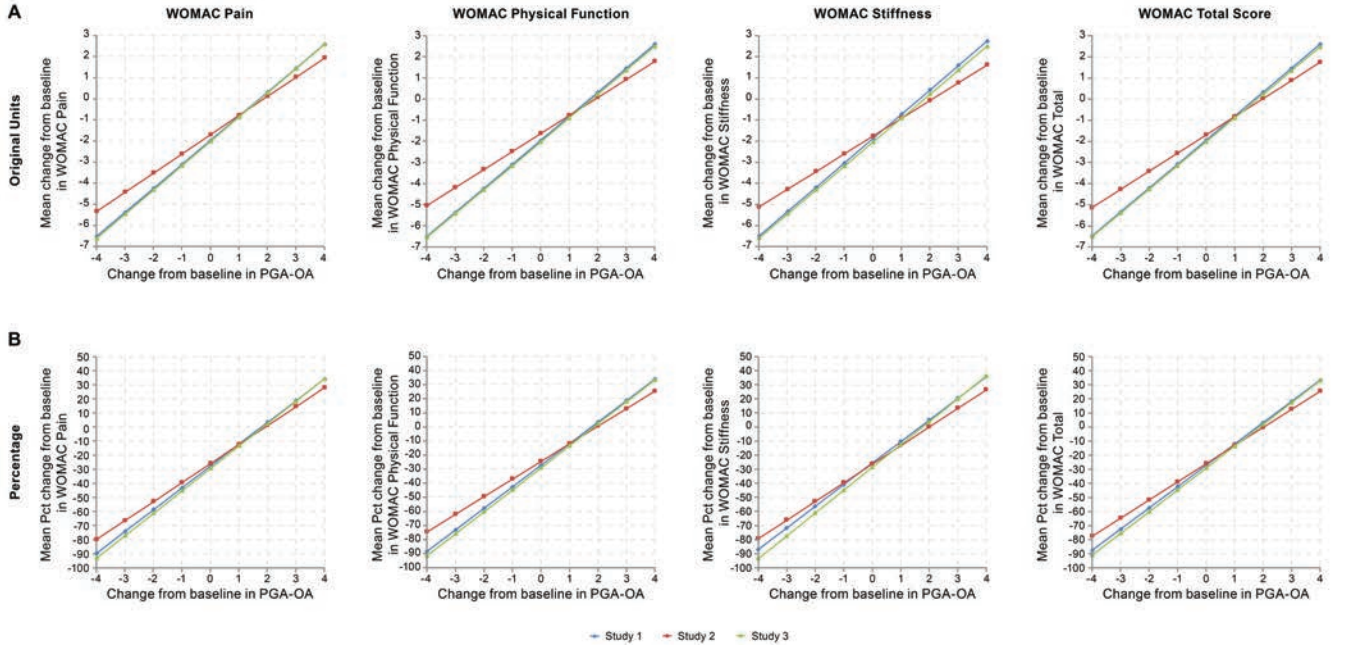

Figure 2. Relationship between change in WOMAC domains and PGA-OA across all studies. The relationship between change in WOMAC domains and PGA-OA was observed across all domains and studies when reported (A) in original units or (B) as a percentage. A negative change represents an improvement in both measures. Pct: percent; PGA-OA: patient global assessment of osteoarthritis; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

category (Figure 3). Although not as clear, the curves for study 2 showed discernable separation, especially for the last 3 of the PGA-OA change categories. This pattern was consistent for all other domains across the 3 studies. In all studies and across all domains, the eCDF curves show that the majority of patients who reported no improvement in PGA-OA saw improvements in WOMAC scores.

DISCUSSION

These analyses of data from 3 randomized clinical trials of patients with moderate-to-severe OA of the hip or knee estimate that a 1-category change on PGA-OA corresponds to WOMAC score changes of 0.84–1.16 (original units) or 12.50–16.23%. A 2-category change on PGA-OA corresponds to an estimated 1.68–2.31 (original units) or 25.01–32.46%

Downloaded on October 10, 2026 from www.jrheum.org

Table 2. Correlations between change in WOMAC domains and change in PGA-OA.

WOMAC Domain	Study 1		Study 2		Study 3	
	Original Units	Percentage	Original Units	Percentage	Original Units	Percentage
Pain	0.613	0.576	0.493	0.477	0.600	0.550
Physical Function	0.624	0.589	0.500	0.491	0.610	0.560
Stiffness	0.549	0.521	0.366	0.358	0.540	0.510
Total	0.615	0.578	0.479	0.479	0.600	0.560

Correlations between change in WOMAC domains and change in PGA-OA at the primary time points of week 16 in studies 1 and 3 and week 24 in study 2. PGA-OA: patient global assessment of osteoarthritis; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

Table 3A. Estimates of meaningful within-patient change by WOMAC domain and study in original units.

WOMAC Domain	Study					
	1		2		3	
	1-category Change in PGA-OA	2-category Change in PGA-OA	1-category Change in PGA-OA	2-category Change in PGA-OA	1-category Change in PGA-OA	2-category Change in PGA-OA
Pain	1.14 (1.08–1.20)	2.28 (2.16–2.40)	0.91 (0.86–0.96)	1.82 (1.72–1.92)	1.15 (1.12–1.18)	2.30 (2.25–2.35)
Physical Function	1.14 (1.08–1.20)	2.28 (2.16–2.40)	0.85 (0.81–0.90)	1.70 (1.61–1.80)	1.13 (1.11–1.16)	2.26 (2.21–2.32)
Stiffness	1.16 (1.09–1.22)	2.31 (2.17–2.45)	0.84 (0.78–0.90)	1.68 (1.56–1.81)	1.14 (1.11–1.17)	2.27 (2.21–2.33)
Total	1.14 (1.08–1.19)	2.27 (2.16–2.39)	0.86 (0.81–0.91)	1.72 (1.63–1.82)	1.12 (1.10–1.15)	2.25 (2.20–2.30)

95% CIs are shown for each value. PGA-OA: patient global assessment of osteoarthritis; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

Table 3B. Estimates of meaningful within-patient change by WOMAC domain and study as percentage change.

WOMAC Domain	Study					
	1		2		3	
	1-category Change in PGA-OA	2-category Change in PGA-OA	1-category Change in PGA-OA	2-category Change in PGA-OA	1-category Change in PGA-OA	2-category Change in PGA-OA
Pain	15.51 (14.69–16.34)	31.02 (29.38–32.67)	13.51 (12.74–14.28)	27.02 (25.48–28.56)	15.93 (15.54–16.32)	31.85 (31.07–32.63)
Physical Function	15.31 (14.51–16.10)	30.61 (29.02–32.21)	12.50 (11.81–13.20)	25.01 (23.61–26.41)	15.63 (15.26–16.01)	31.27 (30.52–32.01)
Stiffness	15.36 (14.42–16.29)	30.72 (28.85–32.59)	13.24 (12.14–14.34)	26.47 (24.28–28.67)	16.23 (15.76–16.70)	32.46 (31.52–33.40)
Total	15.15 (14.37–15.93)	30.30 (28.74–31.85)	12.90 (12.19–13.62)	25.80 (24.37–27.24)	15.51 (15.14–15.88)	31.02 (30.29–31.76)

95% CIs are shown for each value. PGA-OA: patient global assessment of osteoarthritis; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

change in WOMAC, depending on domain and study. The same magnitude of MWPC is applied to meaningful within-person improvement and deterioration, with the 2 differing only in sign. This is supported by evidence of a linear relationship between change in all WOMAC domains and change in PGA-OA.

Estimates for MWPC in WOMAC domains were produced using longitudinal data from all patients, regardless of treatment group or level of change from baseline in the current-state anchor measure of PGA-OA. Consequently, these estimates are applicable at the individual patient level, in accordance with FDA guidance.[7] This methodology contrasts with that used to calculate the MCID, which estimates improvement or worsening in an outcome at the individual or group level by comparing the difference in mean scores between categories of an anchor measure of change, which is subject to recall bias.[1,2,3,8] When calculated at the group level, MCID estimates also need to take the wider disease context into consideration and cannot be used to determine meaningful change in outcomes at the individual patient level.[23]

The estimated values for MWPC produced in this study are similar to and supportive of published data. The definitions of ≥ 10% reduction in pain as minimally important and ≥ 30% reduction as moderately important are consistent with the MWPC in WOMAC for 1- and 2-category changes on

Downloaded on October 10, 2026 from www.jrheum.org

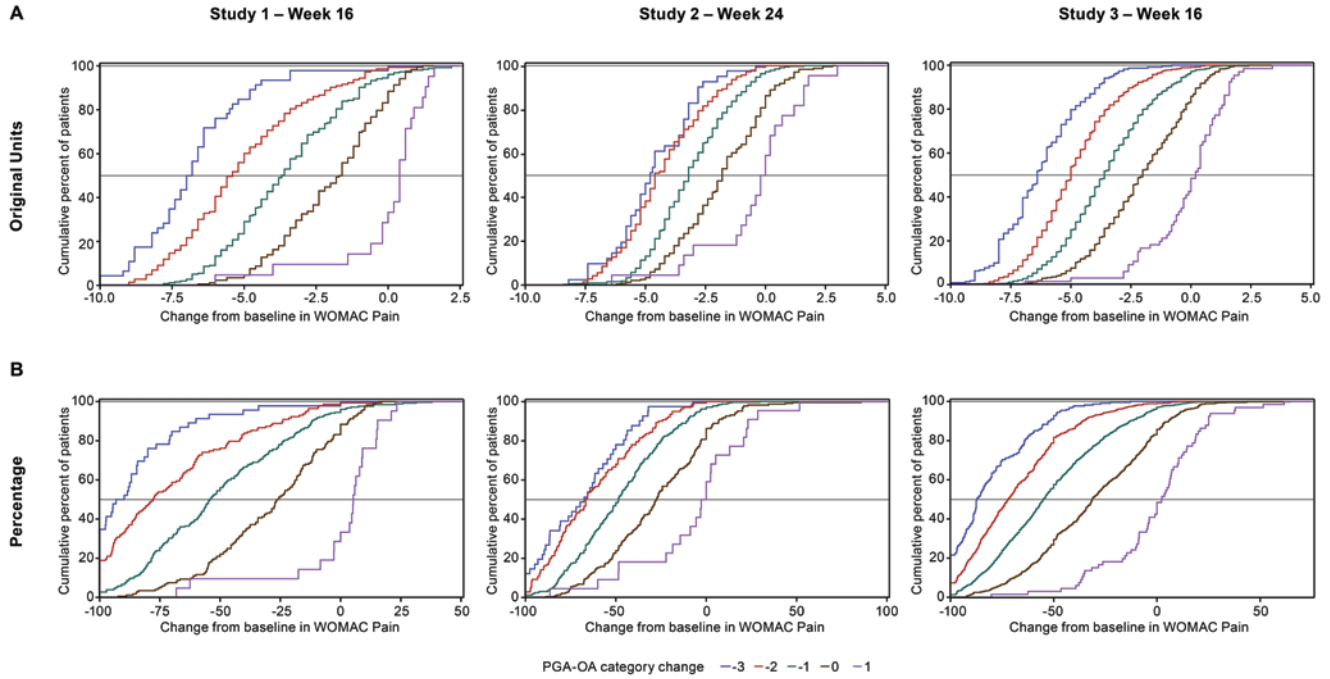

Figure 3. eCDF of change in WOMAC Pain score to support anchor-based models. eCDFs for the change in WOMAC Pain at the primary endpoint of each study, reported (A) in original units or (B) as a percentage. eCDF: empirical cumulative distribution function; PGA-OA: patient global assessment of osteoarthritis; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

PGA-OA, respectively.[4,9] The definitions of an approximately 30% reduction in pain and an approximately 20% reduction in WOMAC Physical Function scores as clinically meaningful are similar to the data presented here.[24] Clinically meaningful changes in patients who received rofecoxib, ibuprofen, or placebo were between 0.9–1.0 points on WOMAC (original 100-mm normalized visual analog scale converted to 0–10 scale), which is similar to the data presented here for a 1-category change on PGA-OA.[25] Finally, the reported values for clinically meaningful changes in WOMAC domains in patients with OA of the hip or knee of between 0.3–1.3 (original standardized 0–100 scale converted to 0–10 scale), depending on domain, are consistent with data for a 1-category change on PGA-OA.[10]

The relationships between change in WOMAC and change in PGA-OA were very similar across all domains and studies. Correlations between changes in WOMAC domains and PGA-OA were approximately 0.5–0.6 for studies 1 and 3, and 0.4–0.5 for study 2. Studies 1 and 3 were primarily based on data from US sites, whereas study 2 was primarily based on European data.

The eCDF curves for all studies show clear separation by changes in PGA-OA category, with the exceptions of –2 and –3 categories in study 2. When interpreting eCDFs and comparing with anchor-based modeling, it should be noted that eCDFs are based on "completers" who had both outcomes collected. eCDFs simply visualize descriptive changes in outcome in a subgroup of subjects at a single timepoint in a study vs our anchor-based model, which uses all available data from all subjects and all timepoints.

Both anchor-based estimates of MWPC and eCDF curves show that patients who reported no change in PGA-OA experienced improvements in WOMAC scores. This could occur because WOMAC and PGA-OA measure similar but distinct concepts. Accounting for this, our model calibrates the relationship between WOMAC and PGA-OA by taking the difference in mean change in WOMAC scores between adjacent categories of the change in PGA-OA anchor; for example, by subtracting the mean change in WOMAC scores between the no-change and 1-category improvement categories on PGA-OA. Such a calibration is analogous to adjusting for placebo in an active intervention study; it is the relative or placebo-adjusted treatment effect that is important, rather than the unadjusted or absolute effect.

The model calibrates the relationship between change in WOMAC and change in PGA-OA over the entire span of their empirically predicted relationship. By contrast, a noncalibrated approach makes no such adjustment and therefore does not adjust for the no-change category on the PGA-OA by subtracting out its corresponding mean change score on WOMAC. This noncalibrated approach forces change on PGA-OA to correspond to no change on WOMAC and therefore assumes that perfect harmony exists in the relationship between the 2 measures, which is not the case in our data and in general.

In conclusion, these analyses of 3 tanezumab studies establish MWPCs for WOMAC domains in patients with moderate-to-severe OA of the hip or knee. A 2-category change on PGA-OA corresponds to an estimated 25.0–32.5% change in WOMAC domains, supporting previous studies describing an approximately 30% reduction in WOMAC as moderately clinically meaningful. An estimated 12.5–16.2% change in WOMAC domains corresponded to a 1-category change on

Downloaded on October 10, 2026 from www.jrheum.org

PGA-OA, suggesting that these improvements may be minimally clinically meaningful.

The large sample sizes used and the reproducibility of estimates for MWPC between each study demonstrate the robustness of the analyses performed. The estimates for MWPC were produced by combining data for both placebo- and tanezumab-treated patients for each study; however, the results are consistent with those of other studies of meaningful change in WOMAC, suggesting that they may be generalizable to this patient population. A possible limitation of this study is the use of only 1 anchor measure, as a credible alternative would have allowed sensitivity analyses to be performed.

These data may aid the interpretation of clinical trials in patients with OA of the hip or knee by defining the meaningful response to treatment in this patient population. Defining MWPC for patients with OA may also be valuable to healthcare professionals, as a means to assess the effect of their interventions. Further work should examine the most appropriate anchor measures and the best methods to determine MWPC.

ACKNOWLEDGMENT

Medical writing support was provided by Steven Moore, PhD, of Engage Scientific Solutions and was funded by Pfizer and Eli Lilly and Company.

DATA SHARING POLICY

Upon request, and subject to review, Pfizer will provide the data that support the findings of this study. Subject to certain criteria, conditions and exceptions, Pfizer may also provide access to the related individual deidentified participant data. See https://www.pfizer.com/science/clinical-trials/trial-data-and-results for more information.

ONLINE SUPPLEMENT

Supplementary material accompanies the online version of this article.

REFERENCES

1. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials 1989;10:407-15.
2. Angst F, Aeschlimann A, Angst J. The minimal clinically important difference raised the significance of outcome effects above the statistical level, with methodological implications for future studies. J Clin Epidemiol 2017;82:128-36.
3. Redelmeier DA, Lorig K. Assessing the clinical importance of symptomatic improvements. An illustration in rheumatology. Arch Intern Med 1993;153:1337-42.
4. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 2008;9:105-21.
5. Dworkin RH, Turk DC, McDermott MP, et al. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. Pain 2009;146:238-44.
6. Guyatt GH, Juniper EF, Walter SD, Griffith LE, Goldstein RS. Interpreting treatment effects in randomised trials. BMJ 1998;316:690-3.
7. U.S. Food and Drug Administration. Patient-focused drug development guidance series for enhancing the incorporation of the patient's voice in medical product development and regulatory decision making. Draft guidance documents. [Internet. Accessed February 24, 2022.] Available from: https://www.fda.gov/drugs/development-approval-process-drugs/fda-patient-focused-drug-development-guidance-series-enhancing-incorporation-patients-voice-medical
8. Schmitt J, Di Fabio RP. The validity of prospective and retrospective global change criterion measures. Arch Phys Med Rehabil 2005;86:2270-6.
9. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001;94:149-58.
10. Doganay Erdogan B, Leung YY, Pohl C, Tennant A, Conaghan PG. Minimal clinically important difference as applied in rheumatology: an OMERACT Rasch Working Group systematic review and critique. J Rheumatol 2016;43:194-202.
11. Ward JR, Williams HJ, Egger MJ, et al. Comparison of auranofin, gold sodium thiomalate, and placebo in the treatment of rheumatoid arthritis. A controlled clinical trial. Arthritis Rheum 1983; 26:1303-15.
12. Deandrade JR, Casagrande PA. A seven-day variability study of 499 patients with peripheral rheumatoid arthritis. Arthritis Rheum 1965;8:302-34.
13. Schnitzer TJ, Easton R, Pang S, et al. Effect of tanezumab on joint pain, physical function, and patient global assessment of osteoarthritis among patients with osteoarthritis of the hip or knee: a randomized clinical trial. JAMA 2019;322:37-48.
14. Berenbaum F, Blanco FJ, Guermazi A, et al. Subcutaneous tanezumab for osteoarthritis of the hip or knee: efficacy and safety results from a 24-week randomised phase III study with a 24-week follow-up period. Ann Rheum Dis 2020;79:800-10.
15. Hochberg MC, Carrino JA, Schnitzer TJ, et al. Long-term safety and efficacy of subcutaneous tanezumab versus nonsteroidal antiinflammatory drugs for hip or knee osteoarthritis: a randomized trial. Arthritis Rheumatol 2021; 73:1167-77.
16. Bellamy N. WOMAC osteoarthritis index: user guide IX. Brisbane: Nicholas Bellamy; 2008.
17. Cappelleri JC, Zou KH, Bushmakin AG, Alvir JMJ, Alemayehu D, Symonds T. Patient-reported outcomes: measurement, implementation and interpretation. New York: Chapman & Hall/CRC; 2013.
18. Fairclough DL. Design and analysis of quality of life studies in clinical trials. New York: Chapman & Hall/CRC; 2010.
19. Fitzmaurice GM, Laird NM, Ware JH. Applied longitudinal analysis. Hoboken: John Wiley & Sons; 2011.
20. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol 2003;56:395-407.
21. 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.
22. Cappelleri JC, Bushmakin AG. Interpretation of patient-reported outcomes. Stat Methods Med Res 2014;23:460-83.
23. Dworkin RH. Two very different types of clinical importance. Contemp Clin Trials 2016;46:11.
24. 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. Ann Rheum Dis 2005;64:29-33.
25. Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities Osteoarthritis Index questionnaire and global assessments in patients with osteoarthritis. J Rheumatol 2000;27:2635-41.

Downloaded on October 10, 2026 from www.jrheum.org