Rheumatol Ther (2021) 8:1759–1774
https://doi.org/10.1007/s40744-021-00372-2

ORIGINAL RESEARCH

Single and Composite Endpoints of Within-Patient Improvement in Symptoms: Pooled Tanezumab Data in Patients with Osteoarthritis

Thomas J. Schnitzer · Francis Berenbaum · Philip G. Conaghan · Robert H. Dworkin · Davide Gatti · Ruoyong Yang · Lars Viktrup · Isabelle Davignon · Christine R. West · Kenneth M. Verburg

Received: July 7, 2021 / Accepted: September 3, 2021 / Published online: October 4, 2021

ABSTRACT

Introduction: Combining measures of key core domains (especially pain and function) into a composite endpoint that requires each patient to meet a threshold of improvement for each domain provides information on multiple aspects of osteoarthritis within individual patients. This pooled analysis of two phase 3 studies (NCT02697773, NCT02709486) explored single and composite endpoints for assessing within-patient improvement in knee or hip osteoarthritis symptoms following subcutaneous administration of tanezumab or placebo.

Methods: Endpoints at week 16 included proportions of responders (≥ 30% improvement) in Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) Pain, WOMAC Physical Function, WOMAC Pain/Function composite, and weekly average pain; and patient acceptable symptom state (PASS) composite responders, minimal clinically important improvement (MCII) composite responders,

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40744-021-00372-2.

T. J. Schnitzer (✉)
Departments of Physical Medicine and Rehabilitation, Internal Medicine/Rheumatology, and Anesthesiology, Northwestern University Feinberg School of Medicine, 710 N Lake Shore Drive, Room 1020, Chicago, IL, USA
e-mail: tjs@northwestern.edu

F. Berenbaum
Department of Rheumatology, Sorbonne Université, INSERM CRSA, AP-HP Hôpital Saint-Antoine, Paris, France

P. G. Conaghan
Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds and NIHR Leeds Biomedical Research Centre, Leeds, UK

R. H. Dworkin
Department of Anesthesiology and Perioperative Medicine, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA

D. Gatti
Department of Rheumatology, University of Verona, Verona, Italy

R. Yang
Pfizer Inc, New York, NY, USA

L. Viktrup
Eli Lilly and Company, Indianapolis, IN, USA

I. Davignon · C. R. West · K. M. Verburg
Pfizer Inc, Groton, CT, USA

Outcome Measures in Rheumatology-Osteoarthritis Research Society International (OMERACT-OARSI) responders, and sustained weekly average pain responders.

Results: Pooled population comprised 1545 patients. Of patients who had a $\geq$ 30% improvement in WOMAC Pain and/or WOMAC Physical Function, 88.5% were WOMAC Pain/Function composite responders, 7.0% were WOMAC Pain (but not Function) responders, and 4.4% were WOMAC Function (but not Pain) responders. Of weekly average pain responders, 43.1% were PASS composite responders. Odds ratios (tanezumab 2.5 mg and 5 mg groups, respectively, vs placebo) were 1.75 and 1.86 (WOMAC Pain/Function composite responders), 1.41 and 1.65 (weekly average pain responders), 1.60 and 1.73 (PASS composite responders), 1.52 and 1.68 (MCII composite responders), 1.75 and 1.88 (OMERACT-OARSI responders), and 1.85 and 1.48 (sustained weekly average pain responders). Subgroup analyses suggested a greater magnitude of effect for patients with a knee index joint compared with hip on some endpoints.

Conclusion: Responders on single pain endpoints were in many cases also responders on function or composite endpoints. Separation of tanezumab from placebo was similar and consistent across single and composite endpoints.

Keywords: Functional status; Osteoarthritis; Pain; Patient-reported outcome measures; Tanezumab

Key Summary Points

Why carry out this study?

Clinicians and patients agree that pain and function are the key core domains to be assessed in osteoarthritis clinical trials. Combining measures of key core domains into a composite endpoint that requires each patient to meet a threshold of improvement for each domain provides information on multiple aspects of osteoarthritis within individual patients.

The objective of these analyses was to explore single and composite endpoints for assessing within-patient improvement in symptoms of osteoarthritis, using pooled data from two studies on subcutaneous administration of tanezumab.

What was learned from the study?

Patients who met improvement thresholds on single pain endpoints were in many cases also responders on function or composite endpoints, and separation of tanezumab from placebo was similar and consistent across single and composite endpoints.

Consideration of the use of composite versus single endpoints depends on many factors, all of which need to be carefully considered when designing a clinical trial.

Determining patients meeting thresholds for improvement in multiple key core domains can provide important clinical information, and the perspectives of patients themselves with respect to composites should be considered.

INTRODUCTION

Osteoarthritis (OA) is one of the most common chronic pain conditions, and is associated with a significant impact on day-to-day functioning and considerable disability [1]. Both clinicians and patients agree that pain and function are the key core domains to be assessed in clinical trials [2], and regulatory authorities may recommend that they be prespecified as co-primary endpoints [3]. Patient's global assessment of disease is another core domain [2] that has been included as a co-primary endpoint in clinical trials [4, 5]. However, co-primary endpoints may each be met by different patients, and clinically it can be important to identify patients meeting both pain and function thresholds. An alternative to multiple single endpoints is the use of a validated composite

outcome measure, which may also provide the benefit of avoiding multiplicity in analyses and allow for reduced sample sizes in clinical trials [6, 7].

A statistically significant improvement compared with placebo needs to be accompanied by clinical relevance [8], such that the magnitude of the improvement meets a validated, clinically meaningful threshold. Within-patient reductions in pain of ≥ 30% or ≥ 50% are often used to represent a clinically meaningful effect [9, 10], although thresholds for meaningful improvement in physical function are less well developed [10]. Other endpoints based on measures of the core domains that have been validated to be clinically meaningful include patient acceptable symptom state (PASS; an absolute value beyond which patients consider themselves well) and minimal clinically important improvement (MCII; the smallest change in measurement that signifies an important improvement in a patient's symptom) [11–13].

Combining measures of the core domains into a composite endpoint that requires each patient to meet a threshold of improvement for each domain provides information on multiple aspects of OA within individual patients. The objective of these analyses was to explore single and composite endpoints for assessing within-patient improvement in symptoms. We employed data pooled from two randomized, placebo-controlled studies on subcutaneous administration of tanezumab in patients with moderate-to-severe OA of the knee or hip [4, 5].

METHODS

Study Design

Both randomized, double-blind, placebo-controlled phase 3 studies were of a similar design with study treatment administered subcutaneously every 8 weeks. Study 1 was a dose-titration study (ClinicalTrials.gov NCT02697773) [4, 14] with primary endpoint at week 16 that had three arms: placebo (at baseline and week 8), tanezumab 2.5 mg (at baseline and week 8), or tanezumab 2.5 mg at baseline and tanezumab 5 mg at week 8. In study 2 (NCT02709486) [5, 15] with primary endpoint at week 24, patients received three doses of placebo, tanezumab 2.5 mg, or tanezumab 5 mg (at baseline, week 8, and week 16). The study protocols were approved by the appropriate institutional review board or independent ethics committee at each participating investigational center, and all patients provided written informed consent prior to entering the studies. The studies were conducted in compliance with the Declaration of Helsinki and International Conference on Harmonisation Good Clinical Practice Guidelines.

Key eligibility criteria were radiographically confirmed (Kellgren–Lawrence [KL] [16] grade ≥ 2 in the index joint) moderate-to-severe OA of the knee or hip; Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) [17] (© 1996 Nicholas Bellamy. WOMAC® is a registered trademark of Nicholas Bellamy (CDN, EU, USA) Pain and Physical Function subscale scores ≥ 5 in the index joint, and patient's global assessment of OA (PGA-OA) of "fair", "poor", or "very poor" at baseline; and a documented history that acetaminophen provided insufficient pain relief, that nonsteroidal anti-inflammatory drugs (NSAIDs) provided inadequate pain relief or could not be taken due to intolerance or contraindication, and that either tramadol or opioids provided inadequate pain relief or could not be taken due to intolerance or contraindication (or the patient was unwilling to take opioids). The index joint was the most painful knee or hip at screening that met pain and radiographic eligibility criteria.

The current pooled, exploratory analyses from week 0 to week 16 were based on WOMAC Pain subscale scores, WOMAC Physical Function subscale scores, average pain scores, and PGA-OA scores. Average pain scores were collected daily using an electronic diary and calculated as weekly means, and the other measures were recorded during clinic visits.

Endpoints

The proportion of patients achieving responder criteria for the endpoints defined below was assessed at week 16, unless otherwise specified.

WOMAC Pain responder: a patient experiencing ≥ 30% or ≥ 50% improvement from baseline in WOMAC Pain subscale.

WOMAC Physical Function responder: a patient experiencing ≥ 30% or ≥ 50% improvement from baseline in WOMAC Physical Function subscale. The established thresholds for clinically meaningful improvement in pain (≥ 30% or ≥ 50% improvement) [9, 10] were adopted for physical function in the current analyses.

WOMAC Pain/Function composite responder: a patient experiencing ≥ 30% or ≥ 50% improvement from baseline in both Pain and Physical Function subscales of WOMAC, assessed at weeks 2, 4, 8, 12, and 16.

Weekly average pain responder: a patient achieving ≥ 30% or ≥ 50% improvement from baseline in weekly average pain score, assessed at weeks 1, 2, 4, 8, 12, and 16.

PASS and MCII were originally based on individual thresholds for pain, function, and patient's global assessment of disease; the thresholds for each scale were defined on a 100-mm visual analog scale (VAS) in a 4-week study in patients experiencing pain from OA (≥ 30 mm on 100-mm VAS) and requiring treatment with an NSAID [11, 12]. The published VAS-based thresholds [11, 12] were adapted for the current analyses based on weekly average pain score (assessed on an 11-point numeric rating scale [NRS], 0–10), WOMAC Physical Function score (NRS, 0–10), and PGA-OA score (5-point Likert scale). For weekly average pain score and WOMAC Physical Function score, the published mean VAS thresholds for pain and function, respectively, were extrapolated to NRS equivalent by dividing by 10. For PGA-OA, the published VAS scores [12] when categorized as 10/30/50/70/90 or 0/25/50/75/100 would correspond to "good" and "very good" on the 5-point scale (PASS); the published VAS improvements [11] were considered closest to an improvement of one category on the 5-point scale (MCII). For the current analyses, PASS and MCII were defined as composite endpoints, such that an individual patient must achieve all three thresholds (pain, function, and global assessment of disease).

PASS composite responder: a patient with weekly average pain score ≤ 3.23 for knee or ≤ 3.50 for hip, WOMAC Physical Function score ≤ 3.10 for knee or ≤ 3.44 for hip, and PGA-OA "good" or "very good".

MCII composite responder: a patient with change (improvement) from baseline in weekly average pain (≥ − 1.99 for knee, ≥ − 1.53 for hip), WOMAC Physical Function (≥ − 0.91 for knee, ≥ − 0.79 for hip), and PGA-OA (improvement of at least one category).

Outcome Measures in Rheumatology-Osteoarthritis Research Society International (OMERACT-OARSI) responder: a patient with improvement from baseline of (i) ≥ 50% and ≥ 2 points in either WOMAC Pain or Physical Function scores, or (ii) ≥ 20% and ≥ 1 point in two of WOMAC Pain, Physical Function, or PGA-OA scores [18].

Sustained pain responder: a patient achieving a pain score ≤ 3 (mild pain) at week 4 that was maintained through week 16, assessed separately for each of WOMAC Pain score and weekly average pain score (based on calculated weekly mean values).

Statistical Analyses

All randomized patients who received at least one dose of placebo or tanezumab in either study were included in the current pooled analyses through week 16 (the primary endpoint for the shorter of the two studies). Pooling the data from these two studies provides a large data set for exploratory analyses. Conservatively, data from the study 1 dose-titration arm (tanezumab 2.5 mg at baseline and tanezumab 5 mg at week 8) were pooled with the study 2 tanezumab 5 mg group for analyses.

Between-group differences were analyzed by logistic regression, the models including baseline WOMAC Pain subscale score, baseline daily average pain score, index joint (hip or knee), treatment, and study; the model for WOMAC Pain/Function composite responders

additionally included baseline WOMAC Physical Function subscale score. A mixed baseline/last observation carried forward approach was used to impute missing data, dependent on the reason for the missing value. Subgroup analyses, based on patients with knee index joint versus hip index joint, were conducted using the same logistic regression models, excluding the index joint term, for weekly average pain responders, WOMAC Pain/Function composite responders, and PASS and MCII composite responders.

All analyses were exploratory and post hoc (not prespecified for the individual studies), except WOMAC Pain responders, WOMAC Physical Function responders, and OMERACT-OARSI responders, which were prespecified secondary endpoints in the individual studies. No correction was made for multiple comparisons in these exploratory pooled analyses. SAS software version 9.4 (Cary, North Carolina) was used for all statistical analyses.

RESULTS

Disposition, Demographics, and Baseline Characteristics

The pooled population comprised 1545 patients. The index joint was a knee for 83.9–84.4% of patients, KL grade 3 for 43.0–45.1% of patients, and KL grade 4 for 32.9–33.5% of patients across the pooled groups (Table 1). Discontinuations from treatment occurred in 9.5–16.3% of patients (Table S1 in the electronic supplementary material [ESM]).

Overall Pooled Population

Of 1039 patients across all pooled treatment groups who had a ≥ 30% improvement in WOMAC Pain and/or WOMAC Physical Function at week 16, 88.5% were WOMAC Pain/Function composite responders, while 7.0% were WOMAC Pain responders but not WOMAC Physical Function responders, and 4.4% were WOMAC Physical Function responders but not WOMAC Pain responders (Fig. 1a). Of 772 patients who had a ≥ 50% improvement in WOMAC Pain and/or WOMAC Physical Function at week 16, 81.6% were WOMAC Pain/Function composite responders, while 12.0% were WOMAC Pain responders but not WOMAC Physical Function responders, and 6.3% were WOMAC Physical Function responders but not WOMAC Pain responders (Fig. 1b).

Concordance between other endpoints varied (Fig. 2). Of weekly average pain responders (≥ 30% improvement), 43.1% of 867 patients met the criteria for PASS composite responder, 77.6% of 865 met the criteria for MCII composite responder, and 95.6% of 865 met the criteria for OMERACT-OARSI response (Fig. 2).

Odds ratios (tanezumab 2.5 mg and 5 mg groups, respectively, vs placebo) for WOMAC Pain/Function composite responders were 1.75 and 1.86 (≥ 30% criterion) and 1.82 and 1.95 (≥ 50% criterion) at week 16 (Fig. 3, Table S2 in the ESM), with consistent and statistically significant separation from placebo at all time points from week 2 (Fig. 4). Odds ratios for weekly average pain responders were 1.41 and 1.65 (≥ 30% criterion) and 1.58 and 1.66 (≥ 50% criterion) at week 16 (Fig. 3, Table S2 in the ESM), with consistent and statistically significant separation from placebo at all time points from week 1 (with the exception of the tanezumab 5 mg group on the ≥ 50% criterion at week 1) (Fig. S1 in the ESM). Odds ratios for PASS composite responders, MCII composite responders, and OMERACT-OARSI responders at week 16 were 1.60 and 1.73, 1.52 and 1.68, and 1.75 and 1.88, respectively (Fig. 3, Table S2 in the ESM). Odds ratios for sustained pain responders were 2.03 and 2.41 (based on WOMAC Pain scores) and 1.85 and 1.48 (based on weekly average pain scores) (Fig. 3, Table S2 in the ESM). Across endpoints, the 95% confidence intervals showed considerable overlap (Fig. 3).

Knee and Hip Subgroups

For the subgroup of patients with a knee as the index joint, odds ratios (tanezumab 2.5 mg and 5 mg groups, respectively, vs placebo) at week 16 were 1.86 and 2.04 (WOMAC Pain/Function composite responders, ≥ 30%

Table 1 Demographics and baseline characteristics of the pooled population

	Placebo, *n* = 514	Tanezumab 2.5 mg, *n* = 514	Tanezumab 5 mg, *n* = 517
Male, *n* (%)	161 (31.3)	171 (33.3)	173 (33.5)
Female, *n* (%)	353 (68.7)	343 (66.7)	344 (66.5)
Age, years, mean (SD)	62.5 (9.8)	63.2 (9.4)	63.4 (9.9)
White/Black or African American/Asian/other or unknown, *n*/*n*	403/60/47/4	423/43/43/5	418/50/42/7
Disease duration, years, mean (SD)[a]	8.7 (8.1)	7.9 (7.8)	8.3 (7.2)
Index joint, *n* (%)			
Hip	80 (15.6)	83 (16.1)	83 (16.1)
Knee	434 (84.4)	431 (83.9)	434 (83.9)
Kellgren–Lawrence grade of index joint, *n* (%)[b]			
0	0	2 (0.4)	0
1	0	1 (0.2)	0
2	124 (24.1)	109 (21.2)	117 (22.7)
3	221 (43.0)	232 (45.1)	226 (43.8)
4	169 (32.9)	170 (33.1)	173 (33.5)
Weekly average pain score, mean (SD)[c]	7.01 (1.48)	6.97 (1.50)	7.00 (1.46)
WOMAC Pain score, mean (SD)[d]	6.9 (1.1)	6.9 (1.1)	6.9 (1.1)
WOMAC Physical Function score, mean (SD)[d]	7.0 (1.1)	7.0 (1.0)	7.0 (1.1)
PGA-OA score, mean (SD)[e]	3.5 (0.6)	3.5 (0.6)	3.5 (0.6)

OA osteoarthritis, *PGA-OA* patient's global assessment of osteoarthritis, *SD* standard deviation, *WOMAC* Western Ontario and McMaster Universities Osteoarthritis Index

[a] Sample size *n* = 514 (placebo), *n* = 512 (tanezumab 2.5 mg), *n* = 515 (tanezumab 5 mg)

[b] Sample size *n* = 514 (placebo), *n* = 514 (tanezumab 2.5 mg), *n* = 516 (tanezumab 5 mg). The index joint was selected as the most painful hip or knee at screening. Kellgren–Lawrence [16] grade for classifying OA severity, ranging from 0 (no OA) to 4 (severe OA)

[c] Sample size *n* = 506 (placebo), *n* = 508 (tanezumab 2.5 mg), *n* = 511 (tanezumab 5 mg). Using an electronic diary, each day during the treatment period, patients assessed their average pain in the index joint during the past 24 h on an 11-point numeric rating scale (0 = no pain, 10 = worst possible pain), in response to: "Select the number that best describes your average pain in your index joint (e.g., the right knee) in the past 24 h". Weekly means were calculated from daily scores

[d] Sample size *n* = 513 (placebo), *n* = 513 (tanezumab 2.5 mg), *n* = 517 (tanezumab 5 mg). At clinic visits, patients completed the Pain and Physical Function subscales of the WOMAC questionnaire, electronically, assessing index joint symptoms within the last 48 h on 11-point numeric rating scales (0 = no pain/difficulty, 10 = extreme pain/difficulty)

[e] The PGA-OA was also assessed during clinic visits on a 5-point Likert scale (1 = very good, 5 = very poor) in response to the question: "Considering all the ways your osteoarthritis in your hip/knee affects you, how are you doing today?"

criterion), 1.91 and 1.99 (WOMAC Pain/Function composite responders, ≥ 50% criterion), 1.49 and 1.79 (weekly average pain responders, ≥ 30% criterion), 1.68 and 1.66 (weekly average pain responders, ≥ 50% criterion), 1.57 and 1.77 (PASS composite

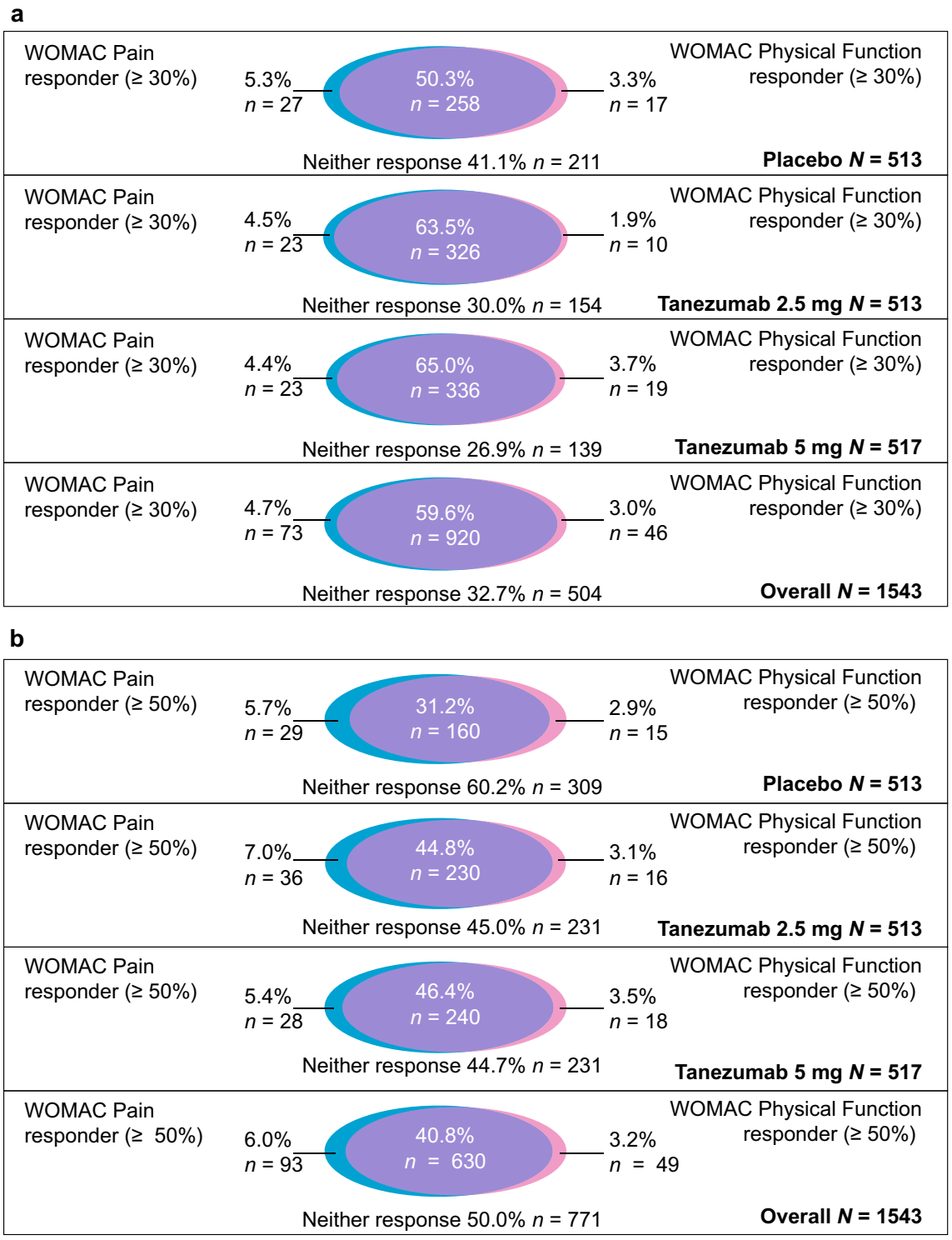
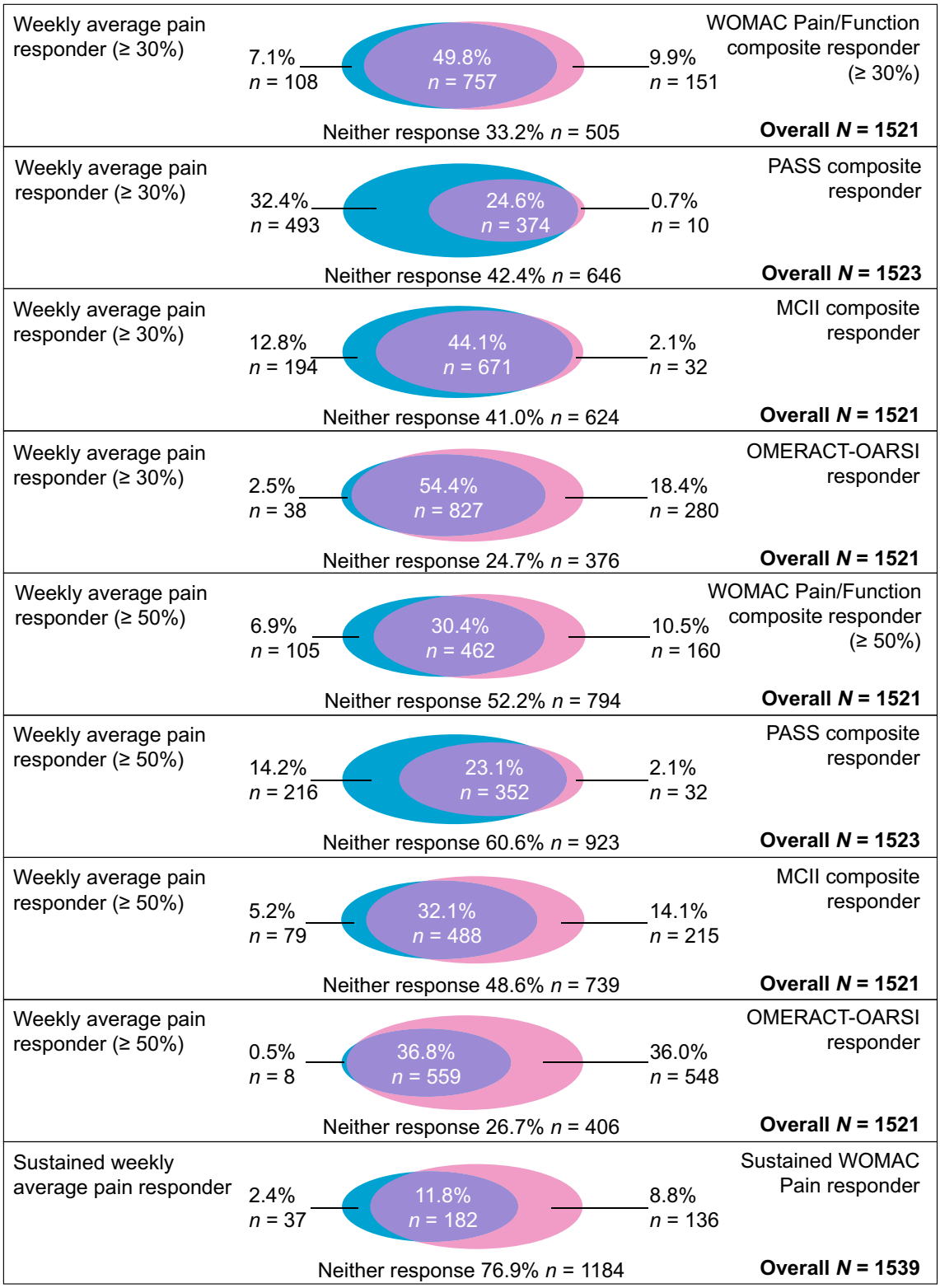


Fig. 1 Venn diagrams of patients achieving responder criteria at week 16 for WOMAC Pain and/or WOMAC Physical Function: **a** ≥ 30% improvement and **b** ≥ 50% improvement (pooled population). Mixed BOCF/LOCF. Denominator for percentages is *N*. *BOCF* baseline observation carried forward, *LOCF* last observation carried forward, *WOMAC* Western Ontario and McMaster Universities Osteoarthritis Index


Weekly average pain responder (≥ 30%)
7.1% n = 108
49.8% n = 757
9.9% n = 151
WOMAC Pain/Function composite responder (≥ 30%)
Neither response 33.2% n = 505
Overall N = 1521
Weekly average pain responder (≥ 30%)
32.4% n = 493
24.6% n = 374
0.7% n = 10
PASS composite responder
Neither response 42.4% n = 646
Overall N = 1523
Weekly average pain responder (≥ 30%)
12.8% n = 194
44.1% n = 671
2.1% n = 32
MCII composite responder
Neither response 41.0% n = 624
Overall N = 1521
Weekly average pain responder (≥ 30%)
2.5% n = 38
54.4% n = 827
18.4% n = 280
OMERACT-OARSI responder
Neither response 24.7% n = 376
Overall N = 1521
Weekly average pain responder (≥ 50%)
6.9% n = 105
30.4% n = 462
10.5% n = 160
WOMAC Pain/Function composite responder (≥ 50%)
Neither response 52.2% n = 794
Overall N = 1521
Weekly average pain responder (≥ 50%)
14.2% n = 216
23.1% n = 352
2.1% n = 32
PASS composite responder
Neither response 60.6% n = 923
Overall N = 1523
Weekly average pain responder (≥ 50%)
5.2% n = 79
32.1% n = 488
14.1% n = 215
MCII composite responder
Neither response 48.6% n = 739
Overall N = 1521
Weekly average pain responder (≥ 50%)
0.5% n = 8
36.8% n = 559
36.0% n = 548
OMERACT-OARSI responder
Neither response 26.7% n = 406
Overall N = 1521
Sustained weekly average pain responder
2.4% n = 37
11.8% n = 182
8.8% n = 136
Sustained WOMAC Pain responder
Neither response 76.9% n = 1184
Overall N = 1539

◀**Fig. 2** Venn diagrams of patients who were responders on both endpoints at week 16 (pooled population). Mixed BOCF/LOCF. Denominator for percentages is *N*. *BOCF* baseline observation carried forward, *LOCF* last observation carried forward, *MCII* minimum clinically important improvement, *OMERACT-OARSI* Outcome Measures in Rheumatology-Osteoarthritis Research Society International, *PASS* patient acceptable symptom state, *WOMAC* Western Ontario and McMaster Universities Osteoarthritis Index

responders), and 1.49 and 1.65 (MCII composite responders) (Fig. 5, Table S3 in the ESM).

For the small subgroup of patients with a hip as the index joint, odds ratios (tanezumab 2.5 mg and 5 mg groups, respectively, vs placebo) at week 16 were 1.32 and 1.19 (WOMAC Pain/Function composite responders, ≥ 30% criterion), 1.44 and 1.78 (WOMAC Pain/Function composite responders, ≥ 50% criterion), 1.08 and 1.32 (weekly average pain responders, ≥ 30% criterion), 1.14 and 1.64 (weekly average pain responders, ≥ 50% criterion), 1.66 and 1.49 (PASS composite responders), and 1.65 and 1.99 (MCII composite responders) (Fig. 5, Table S3 in the ESM).

DISCUSSION

This exploratory analysis of pooled data found that patients who were responders on single pain endpoints were in many cases also responders on function or composite endpoints. Concordance between endpoints varied. Separation of tanezumab treatment effect from placebo was similar and consistent across the endpoints, including single and composite endpoints, with considerable overlap in confidence intervals across endpoints.

The placebo effect is a common factor in studies of OA and pain [19–21] and a large placebo response was reported for the individual tanezumab studies [4, 5]. When the placebo response is large, demonstrating a treatment effect can be more difficult. The use of a validated composite endpoint may enhance sensitivity and has the potential for reducing sample size requirements [6, 7]. However, endpoints measuring OA symptoms are reported to be highly correlated [22] and the findings of studies investigating the responsiveness of composites have been variable [22, 23].

The endpoints investigated here were all measures of within-patient improvement. Since each individual experiences pain differently, assessing within-patient responses to treatment provides valuable information that complements changes in group mean data [10]. The ≥ 30% and ≥ 50% responder thresholds used across the various endpoints were based on those previously established for moderate (≥ 30%) and substantial (≥ 50%) clinically meaningful within-patient reductions in pain [9, 10]. Thresholds for meaningful improvement in physical function are less well developed [10], however, so the same thresholds (≥ 30%, ≥ 50%) were adopted for physical function in the current analyses. The PASS and MCII composite responder endpoints were based on thresholds adapted from those previously published [11, 12]. The patients included in the current pooled population differed from those in the validation studies, which established the ≥ 30% or ≥ 50% improvement in pain threshold in a large, more diverse patient population (including diabetic neuropathy, postherpetic neuralgia, chronic low back pain, fibromyalgia, and OA) [9, 10] and the PASS and MCII thresholds in patients with less severe OA [11, 12] than the current pooled population. The OMERACT-OARSI responder endpoint is well established [18]. The sustained pain responder endpoints investigated in the current analyses were exploratory: the threshold, though unvalidated, reflected achievement and maintenance of pain scores in the mild range (≤ 3), which is likely to be important to patients considering they started the trial with moderate-to-severe pain, based on the eligibility criteria.

In the current pooled population, a small number of patients met the threshold for WOMAC Pain responder but not WOMAC Physical Function responder, and for WOMAC Physical Function responder but not WOMAC Pain responder. Whereas these patients would contribute to individual pain and function endpoints, they did not meet the criteria for the

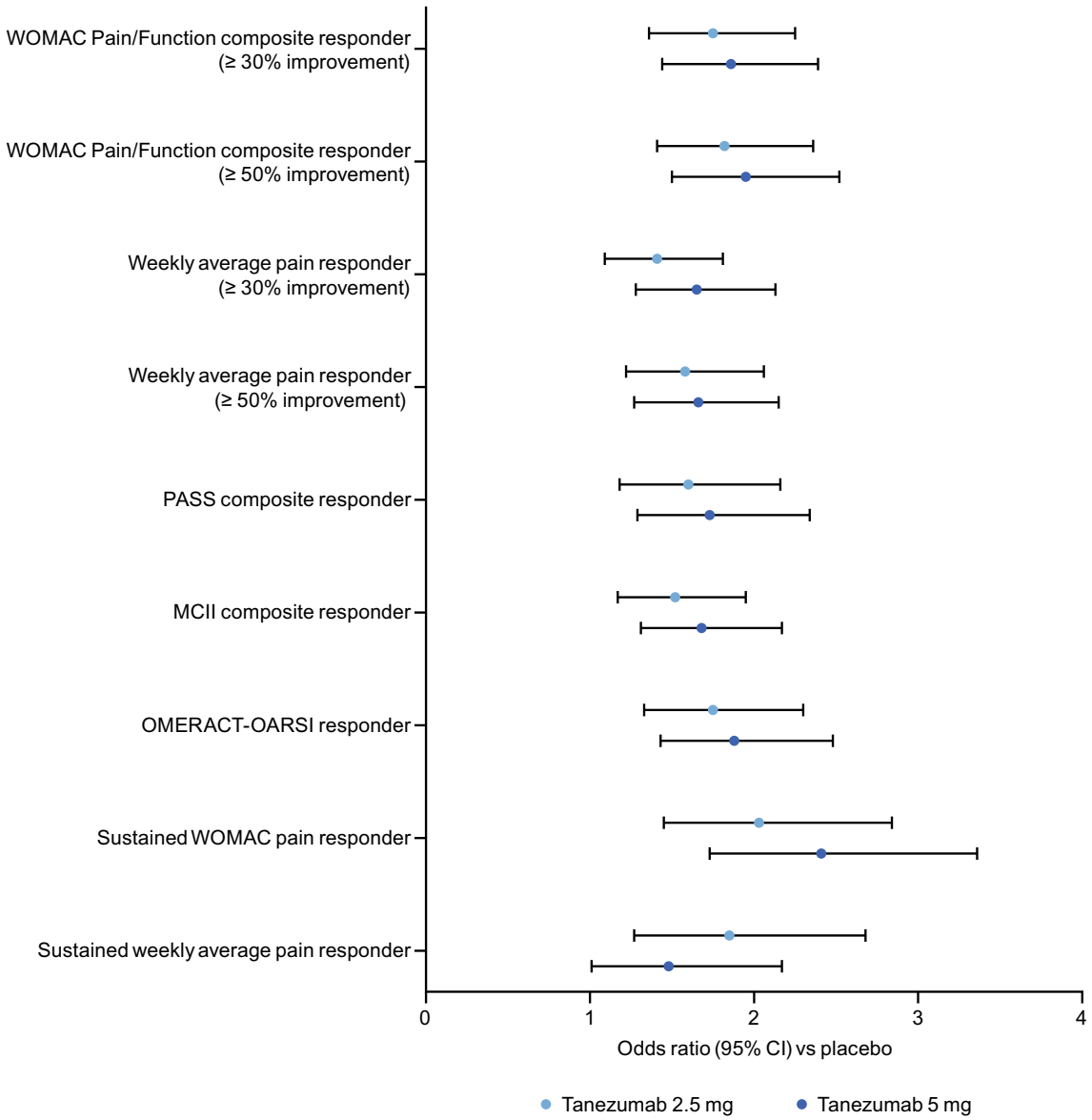


Fig. 3 Separation from placebo at week 16 for the endpoints evaluated (pooled population). Mixed BOCF/LOCF. Logistic regression. See Table S2 in the electronic supplementary material for sample sizes. *BOCF* baseline observation carried forward, *CI* confidence interval, *LOCF* last observation carried forward, *MCII* minimum clinically important improvement, *OMERACT-OARSI* Outcome Measures in Rheumatology-Osteoarthritis Research Society International, *PASS* patient acceptable symptom state, *WOMAC* Western Ontario and McMaster Universities Osteoarthritis Index

composite endpoint (WOMAC Pain/Function composite responders). Hence, the various endpoints can be met by different patients. The PASS and MCII composite responder endpoints added within-patient measures of patient's global assessment of disease to within-patient measures of pain and function, but this did not greatly affect the separation of tanezumab treatment effect from placebo compared with the two-component WOMAC Pain/Function composite responders endpoint.

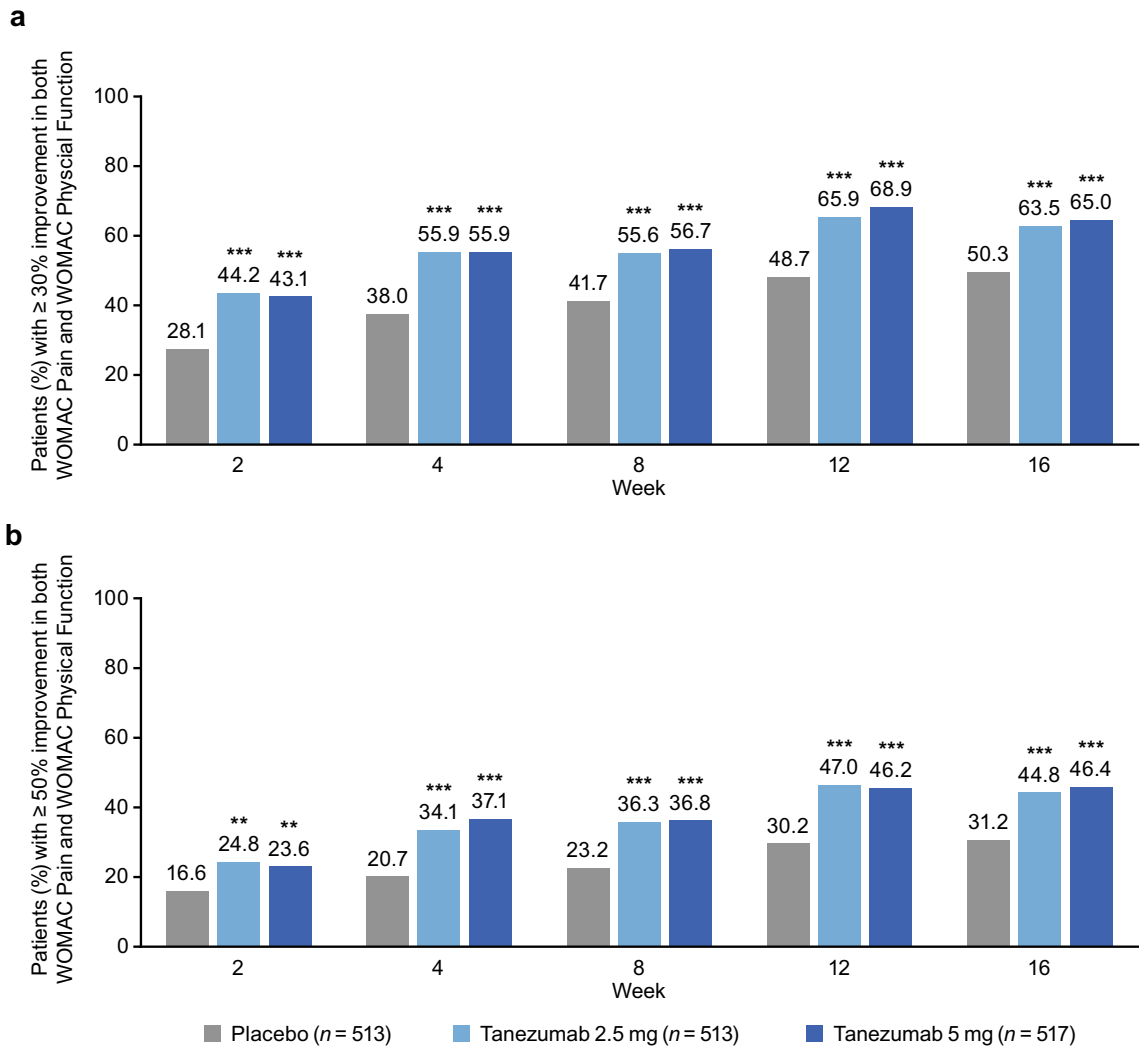


Fig. 4 WOMAC Pain/Function composite responders: proportion of patients achieving **a** ≥ 30% improvement and **b** ≥ 50% improvement, in both WOMAC Pain and WOMAC Physical Function subscales, through week 16 (pooled population). $^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$ versus placebo. Mixed BOCF/LOCF. Logistic regression. *BOCF* baseline observation carried forward, *LOCF* last observation carried forward, *WOMAC* Western Ontario and McMaster Universities Osteoarthritis Index

By considering a longitudinal response that goes beyond discrete time points, the sustained pain responders endpoints take into account the often fluctuating nature of the disease [1]. Of all the endpoints in the current analyses, the greatest separation from placebo (largest odds ratio) was seen with the sustained pain responder endpoint based on WOMAC Pain scores. Interestingly, the odds ratios were lower for sustained pain responders based on weekly average pain scores compared with sustained pain responders based on WOMAC Pain scores. This may be due to the inherent difference in the measures: for weekly average pain, the patient reported daily the average pain in the past 24 h in the electronic diary, while the WOMAC Pain subscale assessed pain during various activities over the 48 h that preceded each clinic visit. Whereas WOMAC Pain score ≥ 5 was part of the eligibility criteria for the studies, weekly average pain score was not.

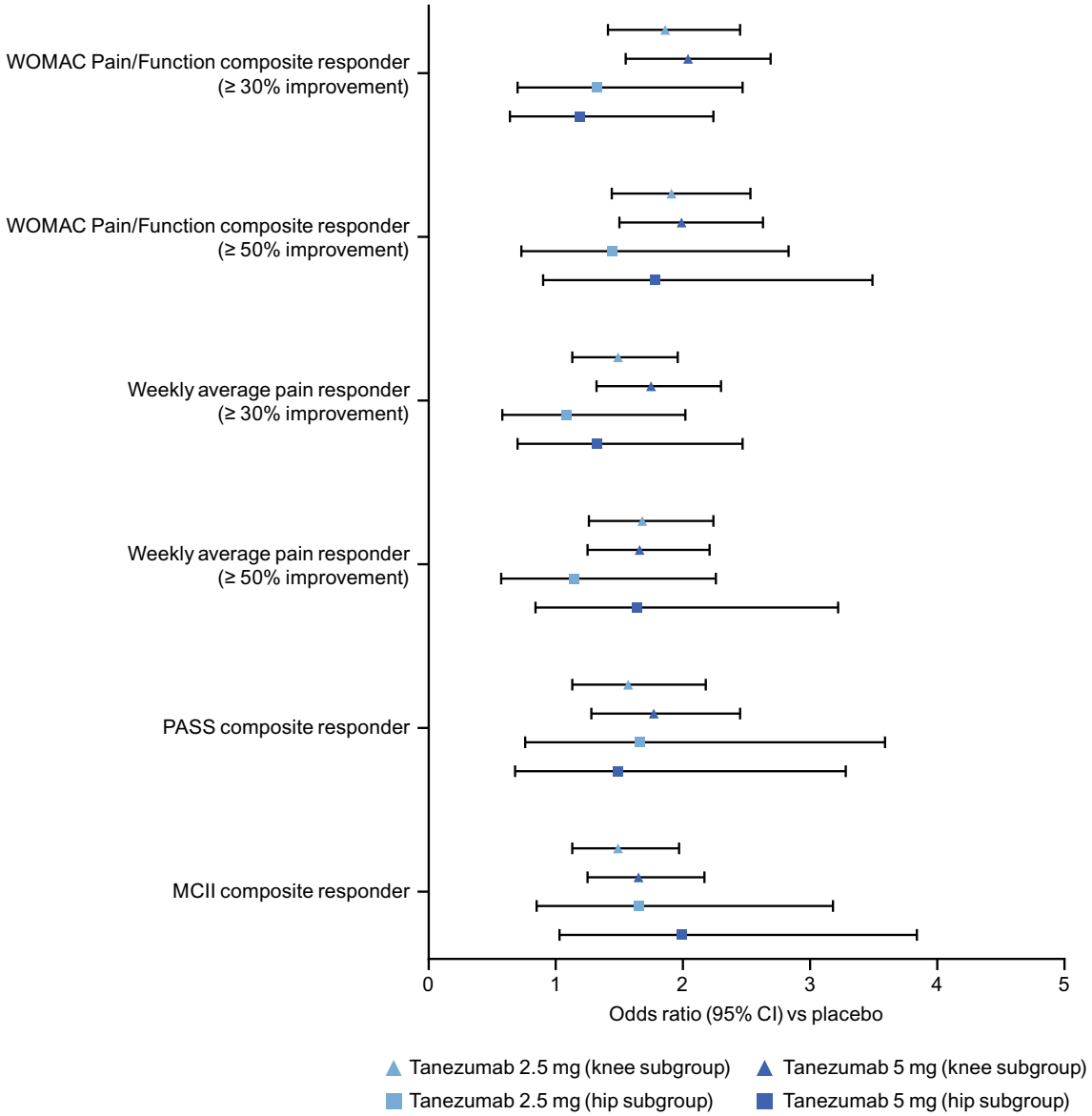


Fig. 5 Separation from placebo at week 16 for the endpoints evaluated (knee and hip subgroups). Mixed BOCF/LOCF. Logistic regression. See Table S3 in the electronic supplementary material. Sample sizes for placebo/tanezumab 2.5 mg/tanezumab 5 mg treatment groups, respectively, 434/430/434 (knee) and 79/83/83 (hip) for WOMAC Pain/Function composite responder; 426/424/431 (knee) and 80/82/80 (hip) for weekly average pain responder; 431/430/433 (knee) and 80/82/83 (hip) for PASS composite responder; and 426/425/431 (knee) and 79/82/80 (hip) for MCII composite responder. *BOCF* baseline observation carried forward, *CI* confidence interval, *LOCF* last observation carried forward, *MCII* minimum clinically important improvement, *PASS* patient acceptable symptom state, *WOMAC* Western Ontario and McMaster Universities Osteoarthritis Index

The separation of tanezumab from placebo seen in the subgroup analyses based on index joint (knee or hip) was in line with that observed for the overall pooled population, although there were some differences between the endpoints. There appeared to be a greater magnitude of treatment effect for patients with a knee index joint compared with a hip index

joint for some endpoints (WOMAC Pain/Function composite responders and pain responders) but not others (PASS and MCII composite responders). These observations need to be interpreted with caution since the numbers of patients in each subgroup were not balanced; there were approximately five times more patients with a knee as the index joint compared with a hip index joint. Despite this limitation, the analyses suggest different composites might be more sensitive in specific circumstances. Previous data were based on intravenous tanezumab administration and did not directly compare patients with knee [24] and hip [25] OA.

The current analyses have limitations. With the exception of WOMAC Pain responders, WOMAC Physical Function responders, and OMERACT-OARSI responders, none of the endpoints were prespecified for the individual studies. The consistency between, and sensitivity of, the various endpoints was not formally tested. The thresholds for clinically meaningful improvement used in the current analyses were adapted from those reported previously, which were based on different patient populations. The hip subgroup was small, and the subgroup analyses should be interpreted cautiously. The current analyses were based on efficacy alone, with analyses up to week 16. Consideration of longer-term efficacy and safety findings [26] is necessary for a full risk–benefit analysis.

CONCLUSION

On the basis of separation from placebo, the single and composite endpoints provided similarly useful information. However, concordance between endpoints varied, and endpoints could be met by different patients. Consideration of the use of composite versus single endpoints depends on many factors, all of which need to be carefully considered when designing a clinical trial. Determining patients meeting thresholds for improvement in multiple key core domains can provide important clinical information, and the perspectives of patients themselves with respect to composites should be considered. The endpoints in the current analysis all demonstrated similar treatment effects for tanezumab compared with placebo in terms of clinically meaningful within-patient improvements in pain, and within-patient improvements in function and composites based on within-patient measures of pain, function, and patient's global assessment of disease.

ACKNOWLEDGEMENTS

Philip G. Conaghan is supported in part through the United Kingdom National Institute for Health Research (NIHR) Leeds Biomedical Research Centre (BRC). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Funding. The study was sponsored by Pfizer and Eli Lilly and Company. Publication fees (expenses to the journal for Rapid Service and Open Access) were funded by Pfizer and Eli Lilly and Company. Pfizer is the manufacturer of tanezumab, which is being investigated for the treatment of patients with chronic pain. Manuscript authors from Pfizer contributed to the study design; data collection, management, and interpretation of data; and the preparation, review, and approval of the manuscript. Manuscript authors from Eli Lilly and Company contributed to the study design; interpretation of data; and the preparation, review, and approval of the manuscript.

Medical Writing Assistance. Medical writing support was provided by Kim Russell, PhD, of Engage Scientific Solutions (Horsham, UK) and was funded by Pfizer and Eli Lilly and Company.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Author Contributions. All authors contributed to the conception or design of the study. Ruoyong Yang, Isabelle Davignon, Christine R. West, and Kenneth M. Verburg contributed to the acquisition of data. All authors contributed to the analysis or interpretation of data. All authors contributed to drafting the manuscript and revising it critically for important intellectual content. All authors approved the final version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Disclosures. Thomas J. Schnitzer reports clinical research study support (Pfizer, Lilly, Regeneron, Galapagos, Taiwan Liposome Corporation, Anika Therapeutics) and fees for consultancy/advisory boards (Pfizer, Lilly, GSK, AstraZeneca, Galapagos, Merck). Francis Berenbaum reports grant/research support (TRB Chemedica) and fees for consultancy (Boehringer, Bone Therapeutics, CellProthera, Expanscience, Galapagos, Gilead, GSK, Merck Sereno, MSD, Nordic, Novartis, Pfizer, Regulaxis, Roche, Sandoz, Sanofi, Servier, UCB, Peptinov, 4P Pharma, 4Moving Biotech). Philip G. Conaghan reports fees for speaker's bureaus (AbbVie, Novartis) and consultancy (AstraZeneca, BMS, Eli Lilly, EMD Serono, Flexion Therapeutics, Galapagos, Gilead, Novartis, Pfizer). Robert H. Dworkin has received in the past 5 years research grants and contracts (US Food and Drug Administration and the US National Institutes of Health), and compensation for serving on advisory boards or consulting on clinical trial methods (Abide, Acadia, Adynxx, Analgesic Solutions, Aptinyx, Aquinox, Asahi Kasei, Astellas, AstraZeneca, Biogen, Biohaven, Boston Scientific, Braeburn, Cardialen, Celgene, Centrexion, Chromocell, Clexio, Collegium, Concert, Confo, Decibel, Dong-A, Editas, Eli Lilly, Ethismos [equity], Eupraxia, Glenmark, Grace, Hope, Immune, Lotus, Mainstay, Merck, Neumentum, Neurana, NeuroBo, Novaremed, Novartis, Olatec, Pfizer, Phosphagenics, Quark, Reckitt Benckiser, Regenacy [also equity], Relmada, Sanifit, Scilex, Semnur, SIMR Bio, SK Life Sciences, Sollis, SPRIM, Teva, Theranexus, Trevena, Vertex, and Vizuri). Davide Gatti reports fees for speaker's bureaus (Celgene, Neopharmed-Gentili, Eli Lilly, UCB) and consultancy (Abiogen, MSD Italia, Pfizer). Ruoyong Yang, Isabelle Davignon, Christine R. West, and Kenneth M. Verburg are employees of Pfizer with stock and/or stock options. Lars Viktrup is an employee of Eli Lilly and Company and owns stock in Lilly.

Prior Presentation. Some of these data were presented at American College of Rheumatology (ACR) Convergence, November 5–9, 2020.

Compliance with Ethics Guidelines. The study protocols were approved by the appropriate institutional review board or independent ethics committee at each participating investigational center (main IRB for Study 1: Schulman Associates IRB/Advarra, Ohio, United States; for Study 2, see supplementary Appendix 1). All patients provided written informed consent prior to entering the studies. The studies were conducted in compliance with the Declaration of Helsinki and International Conference on Harmonisation Good Clinical Practice Guidelines.

Data Availability. Upon request, and subject to certain criteria, conditions, and exceptions (see https://www.pfizer.com/science/clinical-trials/trial-data-and-results for more information), Pfizer will provide access to individual de-identified participant data from Pfizer-sponsored global interventional clinical studies conducted for medicines, vaccines, and medical devices (1) for indications that have been approved in the US and/or EU or (2) in programs that have been terminated (i.e., development for all indications has been discontinued). Pfizer will also consider requests for the protocol, data dictionary, and statistical analysis plan. Data may be requested from Pfizer trials 24 months after study completion. The de-identified participant data will be made available to researchers whose proposals meet the research criteria and other conditions, and for which an exception does not apply, via a secure portal. To gain access, data requestors must enter into a data access agreement with Pfizer.

REFERENCES

1. Osteoarthritis Research Society International. White paper: osteoarthritis: a serious disease. Submitted to the US Food and Drug Administration December 1, 2016. https://www.oarsi.org/sites/default/files/docs/2016/oarsi_white_paper_oa_serious_disease_121416_1.pdf. Accessed 1 Sep 2020.

2. Smith TO, Hawker GA, Hunter DJ, et al. The OMERACT-OARSI core domain set for measurement in clinical trials of hip and/or knee osteoarthritis. J Rheumatol. 2019;46:981–9.

3. European Medicines Agency. Guideline on clinical investigation of medicinal products used in the treatment of osteoarthritis. 2010. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-used-treatment-osteoarthritis_en.pdf. Accessed 11 Nov 2020.

4. Schnitzer TJ, Easton R, Pang S, et al. Efficacy and safety of tanezumab for the treatment of osteoarthritis pain in the hip or knee. A randomized clinical trial. JAMA. 2019;322:37–48.

5. 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 3 study with a 24-week follow-up period. Ann Rheum Dis. 2020;79:800–10.

6. Turk DC, Dworkin RH, McDermott MP, et al. Analyzing multiple endpoints in clinical trials of pain treatments: IMMPACT recommendations. Initiative on methods, measurement, and pain assessment in clinical trials. Pain. 2008;139:485–93.

7. Food and Drug Administration. Multiple endpoints in clinical trials: guidance for industry. Draft guidance January 2017. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/multiple-endpoints-clinical-trials-guidance-industry. Accessed 7 Dec 2020.

8. Wasserstein RL, Lazar NA. The ASA statement on *p* values: context, process, and purpose. Am Stat. 2016;70:129–33.

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. 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.

11. 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.

12. 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. Ann Rheum Dis. 2005;64:34–7.

13. Bellamy N, Hochberg M, Tubach F, et al. Development of multinational definitions of minimal clinically important improvement and patient acceptable symptomatic state in osteoarthritis. Arthritis Care Res (Hoboken). 2015;67:972–80.

14. Schnitzer TJ, Khan A, Bessette L, et al. Onset and maintenance of efficacy of subcutaneous tanezumab in patients with moderate to severe osteoarthritis of the knee or hip: a 16-week dose-titration study. Semin Arthritis Rheum. 2020;50:387–93.

15. Berenbaum F, Langford R, Perrot S, et al. Subcutaneous tanezumab for osteoarthritis: is the early improvement in pain and function meaningful and sustained over 24 weeks? Eur J Pain. 2021;25:1525–39.

16. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957;16:494–502.

17. Theiler R, Spielberger J, Bischoff HA, Bellamy N, Huber J, Kroesen S. Clinical evaluation of the WOMAC 3.0 OA Index in numeric rating scale format using a computerized touch screen version. Osteoarthr Cartil. 2002;10:479–81.

18. Pham T, van der Heijde D, Altman RD, et al. OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited. Osteoarthr Cartil. 2004;12:389–99.

19. Zhang W, Robertson J, Jones AC, Dieppe PA, Doherty M. The placebo effect and its determinants in osteoarthritis: meta-analysis of randomised controlled trials. Ann Rheum Dis. 2008;67:1716–23.

20. Greene CS, Goddard G, Macaluso GM, Mauro G. Topical review: placebo responses and therapeutic responses. How are they related? J Orofac Pain. 2009;23:93–107.

21. Zou K, Wong J, Abdullah N, et al. Examination of overall treatment effect and the proportion attributable to contextual effect in osteoarthritis: meta-analysis of randomised controlled trials. Ann Rheum Dis. 2016;75:1964–70.

22. Bolognese JA, Ehrich EW, Schnitzer TJ. Precision of composite measures of osteoarthritis efficacy in comparison to that of individual endpoints. J Rheumatol. 2001;28:2700–4.

23. Trudeau J, Van Inwegen R, Eaton T, et al. Assessment of pain and activity using an electronic pain diary and actigraphy device in a randomized, placebo-controlled crossover trial of celecoxib in osteoarthritis of the knee. Pain Pract. 2015;15:247–55.

24. Brown MT, Murphy FT, Radin DM, Davignon I, Smith MD, West CR. Tanezumab reduces osteoarthritic knee pain: results of a randomized, double-blind, placebo-controlled phase III trial. J Pain. 2012;13:790–8.

25. Brown MT, Murphy FT, Radin DM, Davignon I, Smith MD, West CR. Tanezumab reduces osteoarthritic hip pain: results of a randomized, double-blind, placebo-controlled phase III trial. Arthritis Rheum. 2013;65:1795–803.

26. 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.