Time to first and sustained improvement in WOMAC domains among patients with osteoarthritis receiving tanezumab

D.J. Hunter,*, T.J. Schnitzer, J. Hall, D. Semel, I. Davignon, J.C. Cappelleri, A.G. Bushmakine, L. Abraham

a University of Sydney, Sydney, Australia
b Northwestern University Feinberg School of Medicine, Chicago, IL, USA
c Eli Lilly and Company, Indianapolis, IN, USA
d Pfizer Inc, New York, NY, USA
e Pfizer Inc, Groton, CT, USA
f Pfizer Ltd, Tadworth, UK

ABSTRACT

Objective: To assess onset of effect in three placebo- or nonsteroidal anti-inflammatory drug (NSAID)-controlled trials of tanezumab in patients with moderate-to-severe osteoarthritis.

Methods: Post-hoc nonparametric Kaplan–Meier analyses were used to estimate median time to first improvement and to sustained improvement in Western Ontario and McMaster Universities Osteoarthritis Index domain (Pain, Physical Function, Stiffness) scores across a range of improvement thresholds (0–100%, in 5% increments). Time to first improvement was defined as the first week scores met the pre-specified threshold. Time to sustained improvement was defined as the first week scores met the pre-specified threshold and were sustained (on average) for the remainder of the treatment period.

Results: Across all domains, tanezumab-treated patients had shorter median times to first improvement (at most thresholds) and reached higher levels of improvement than placebo-treated patients. No substantial differences were observed between tanezumab doses (2.5 and 5 mg), or between tanezumab and NSAIDs. Most patients experiencing an event of first improvement went on to experience a sustained event. At low thresholds, sustained improvement occurred simultaneously with, or shortly after, first improvement. At higher thresholds, median time to sustained improvement was longer than median time to first improvement.

Conclusions: Following initiation of tanezumab treatment, first improvement of osteoarthritis symptoms of 30% was evident within 2–4 weeks and sustained improvement was evident within 2–8 weeks. Time to improvement of 50% was more variable, with first and sustained events expected within 4–16 and 8–24 weeks, respectively.

ClinicalTrials.gov identifiers: NCT02697773; NCT02709486; NCT02528188.

1. Introduction

Healthcare is evolving from a model where treatment decisions are primarily driven by the physician to a more collaborative model where patients share responsibility and are more actively involved in the decision-making process [1–3]. A key factor in this collaborative model is setting realistic treatment expectations, since patient expectation may impact adherence, satisfaction, and outcomes in chronic conditions [4–8]. Knowledge of an intervention’s onset, extent, and duration of effect helps set realistic treatment expectations. An understanding of onset of effect establishes a timeframe in which patients can expect to experience a treatment benefit and can help prevent unnecessarily prolonged exposure. An understanding of extent of effect establishes how much benefit a patient can generally expect in response to treatment. Finally, an understanding of duration of effect establishes whether a patient can expect to maintain treatment benefits and, if so, for how long.

Osteoarthritis (OA) is a chronic condition where, in the absence of disease-modifying agents, treatment focuses on alleviation of symptoms, particularly pain and loss of physical function [9,10]. Tanezumab, a monoclonal antibody against nerve growth factor (NGF), has demonstrated an ability to improve pain and physical function in patients with moderate to severe OA and a history of inadequate response or
intolerance to standard-of-care analgesics [11–14]. Inhibition of NGF has both short-term and long-term effects on nociceptive signaling and represents a novel mechanism of action in the context of OA treatment [15]. While tanezumab’s onset of effect is not expected to be as immediate as analgesics with fast-acting mechanisms of actions, effects are expected to be maintained throughout the dosing interval (8 weeks) and over long-term treatment durations [14,16]. Here, we utilize a novel series of nonparametric models to estimate median times to onset of treatment.

Table 1
Patient demographics and clinical characteristics.

| Study 1 | PBO | TNZ 2.5 mg | TNZ 2.5/5 mg |
|--------|-----|------------|-------------|
| (N = 232) | (N = 231) | (N = 233) |
| Gender, n (%) | | | |
| Male | 75 (32.3) | 86 (37.2) | 82 (35.2) |
| Female | 157 (67.7) | 145 (62.8) | 151 (64.8) |
| Race | | | |
| White | 156 (67.2) | 178 (77.1) | 170 (73.0) |
| Black or African American | 60 (25.9) | 45 (18.6) | 50 (21.5) |
| Asian | 13 (5.6) | 5 (2.2) | 8 (3.4) |
| Other | 3 (1.3) | 5 (2.2) | 5 (2.1) |
| Mean (SD) age, years | 60.4 (9.8) | 60.9 (10.0) | 61.2 (9.0) |
| Index joint, n (%) | | | |
| Knee | 199 (85.8) | 197 (85.3) | 198 (85.0) |
| Hip | 33 (14.2) | 34 (14.7) | 35 (15.0) |
| Mean (SD) WOMAC Pain score at baseline | 7.30 (1.15) | 7.08 (1.16) | 7.33 (1.26) |
| Mean (SD) WOMAC Physical Function score at baseline | 7.18 (1.11) | 7.18 (1.11) | 7.18 (1.18) |
| Mean (SD) WOMAC Stiffness score at baseline | 7.49 (1.38) | 7.27 (1.40) | 7.58 (1.40) |
| Completed the treatment period, n (%) | 192 (82.8) | 208 (90.0) | 209 (89.7) |

| Study 2 | PBO | TNZ 2.5 mg | TNZ 5 mg |
|--------|-----|------------|--------|
| (N = 282) | (N = 283) | (N = 284) |
| Gender, n (%) | | | |
| Male | 86 (30.5) | 85 (30.0) | 91 (32.0) |
| Female | 196 (69.5) | 198 (70.0) | 193 (68.0) |
| Race | | | |
| White | 247 (87.6) | 245 (86.6) | 248 (87.3) |
| Black or African American | 0 | 0 | 0 |
| Asian | 34 (12.1) | 38 (13.4) | 34 (12.0) |
| Other | 0 | 0 | 2 (0.7) |
| Mean (SD) age, years | 64.2 (9.6) | 65.2 (8.4) | 65.2 (10.2) |
| Index joint, n (%) | | | |
| Knee | 235 (83.3) | 234 (82.7) | 236 (83.1) |
| Hip | 47 (16.7) | 49 (17.3) | 48 (16.9) |
| Mean (SD) WOMAC Pain score at baseline | 6.59 (0.94) | 6.70 (0.94) | 6.60 (0.89) |
| Mean (SD) WOMAC Physical Function score at baseline | 6.67 (0.87) | 6.77 (0.87) | 6.76 (0.88) |
| Mean (SD) WOMAC Stiffness score at baseline | 6.46 (1.43) | 6.44 (1.59) | 6.44 (1.53) |
| Completed the treatment period, n (%) | 238 (84.4) | 257 (90.8) | 255 (89.8) |

| Study 3 | NSAID | TNZ 2.5 mg | TNZ 5 mg |
|--------|-------|------------|--------|
| (N = 996) | (N = 1002) | (N = 998) |
| Gender, n (%) | | | |
| Male | 334 (33.5) | 365 (36.4) | 344 (34.5) |
| Female | 662 (66.5) | 637 (63.6) | 654 (65.5) |
| Race | | | |
| White | 680 (68.3) | 705 (70.4) | 712 (71.3) |
| Black or African American | 186 (18.7) | 166 (16.6) | 162 (16.2) |
| Asian | 99 (9.9) | 110 (11.0) | 95 (9.5) |
| Other | 31 (3.1) | 21 (2.1) | 29 (2.9) |
| Mean (SD) age, years | 60.3 (9.5) | 60.3 (9.2) | 61.2 (9.6) |
| Index joint, n (%) | | | |
| Knee | 852 (85.5) | 851 (84.9) | 850 (85.2) |
| Hip | 144 (14.5) | 151 (15.1) | 148 (14.8) |
| Mean (SD) WOMAC Pain score at baseline | 6.96 (1.08) | 7.01 (1.12) | 7.02 (1.12) |
| Mean (SD) WOMAC Physical Function score at baseline | 7.09 (1.09) | 7.09 (1.07) | 7.08 (1.11) |
| Mean (SD) WOMAC Stiffness score at baseline | 7.09 (1.42) | 7.15 (1.42) | 7.20 (1.40) |
| Completed the treatment period, n (%) | 446 (44.8) | 447 (44.6) | 419 (42.0) |

Abbreviations: BID = twice daily, NSAID = nonsteroidal anti-inflammatory drug, PBO = placebo, SD = standard deviation, TNZ = tanezumab, WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

a Patients received a 2.5 mg dose at baseline and a 5 mg dose at Week 8.

b NSAID regimen included naproxen 500 mg BID, celecoxib 100 mg BID, or diclofenac extended release 75 mg BID.

c In Study 3, patients had to meet the following criteria in order to continue receiving SC study medication at and beyond Week 16: a ≥30% reduction in WOMAC Pain subscale relative to baseline in the index joint at Week 16 and a ≥15% reduction in WOMAC Pain subscale relative to baseline in the index joint at Week 2, 4, or 8.
effect (across different levels of symptomatic improvement) in three recently completed trials of tanezumab.

On October 26, 2021, Pfizer Inc. and Eli Lilly and Company announced the discontinuation of the tanezumab global clinical development program as a result of the outcomes of regulatory reviews of tanezumab for the treatment of osteoarthritis pain by the U.S. Food and Drug Administration and European Medicines Agency [17,18].

2. Methods

2.1. Data sources

Patient-level data were derived from three randomized double-blind controlled trials of subcutaneous (SC) tanezumab in patients with moderate to severe OA of the knee or hip and a history of inadequate response or inability to take standard-of-care analgesics: NCT02697773 (Study 1), NCT02709486 (Study 2), and NCT02528188 (Study 3) [11,13,14]. Study 1 included 16-week treatment with SC placebo, SC tanezumab 2.5 mg, or SC tanezumab 2.5/5 mg (switch from 2.5 to 5 mg at Week 8) [11]. Study 2 included 24-week treatment with SC placebo, SC tanezumab 2.5 mg, or SC tanezumab 5 mg [13]. Study 3 included 56-week treatment with SC tanezumab 2.5 mg, SC tanezumab 5 mg, or oral nonsteroidal anti-inflammatory drugs (NSAIDs) [14]. SC treatment was administered every 8 weeks, rescue medication (acetaminophen/paracetamol) was allowed, and occasional use of analgesics for self-limiting conditions unrelated to OA was allowed during each study treatment period [11,13,14].

2.2. Endpoints

The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC; © 1996 Nicholas Bellamy; WOMAC® is a registered trademark of Nicholas Bellamy (CDN, EU, USA)) was used to assess symptoms of OA [19]. In each study, change in WOMAC Pain and Physical Function domain scores were co-primary endpoints (along with Patient Global Assessment of OA), while change in WOMAC Stiffness domain score was a secondary endpoint. The WOMAC Pain, Physical Function, and Stiffness domains consist of five, 17, and two questions, respectively, with a recall period of the previous 48 h. Each domain has an overall score that ranges from 0 to 10, with higher scores indicating greater pain intensity, interference with physical function, or level of stiffness, respectively. In addition to baseline, the WOMAC was completed at weeks 2, 4, 8, 12, and 16 for Study 1; weeks 2, 4, 8, 12, 16, and 24 for Study 2; and weeks 2, 4, 8, 16, 24, 32, 40, 48, and 56 for Study 3.

2.3. Time-to-event analyses

A series of post-hoc nonparametric time-to-event (Kaplan–Meier) analyses were used to estimate median time to first improvement and sustained improvement in the WOMAC domains. Improvement thresholds were evaluated in increments of 5% from 5% to 100% (i.e., 5%, 10% ... 95%, 100%). Time to first improvement was defined as the first post-baseline week WOMAC domain scores met the pre-specified threshold. For example, if the first time a subject achieves at least 30% improvement (from baseline) was study week 4, then time to first improvement of 30%
is week 4. Time to sustained improvement was defined as the first post-baseline week at WOMAC domain scores met the pre-specified threshold and were then sustained (on average) for the remainder of the treatment period. If, in our example, the subject achieved at least 30% improvement at week 4 and then also averaged at least 30% improvement until the end of the study (including at week 4), then this is a stable event and time to stable improvement of 30% for this subject is week 4.

In time-to-event analyses, patients who do not achieve an event by the end of study are typically censored [20]. Simply ignoring censored observations, however, is not appropriate. Our analysis, therefore, included both uncensored and censored observations. Three different scenarios defined the time of censor if an event was not achieved during the study (i.e., the threshold of improvement was never met). If a subject had baseline data (not 0) and post-baseline data but did not achieve an event during the study, the subject was censored at the last post-baseline timepoint that had data. If a subject did not have baseline data (or had a score of 0 at baseline) but had post-baseline data, the subject was censored at the last post-baseline timepoint that had data. If a subject had baseline data (not 0) but no post-baseline data, the subject was censored at baseline.

2.4. Statistical analyses

Since time-to-event values are usually not normally distributed, we applied (nonparametric) Kaplan–Meier time-to-event analyses [21,22]. Analyses were stratified by treatment and conducted using SAS Proc Lifetest. This nonparametric analysis included both censored and uncensored observations. It also included a test of equality over strata (using a log-rank test) allowing for comparison of time-to-event curves [23,24]. Significance was declared at the 0.05 level. All comparisons were unadjusted for multiplicity. The 3 trials included in this study were analyzed individually (rather than pooled) since one of trials utilized a different comparator and, more importantly, all of the trials had different durations of treatment.

3. Results

3.1. Patients

Across studies, patients were predominantly white (74%), were female (66%), and had a knee as the index joint (85%) (Table 1). Mean age was ~65 years in Study 2 and ~61 years in Studies 1 and 3. Baseline WOMAC scores were similar across treatment arms in each study. Over 80% of patients completed treatment in Studies 1 (16 weeks) and 2 (24 weeks), whereas just over 40% completed treatment in Study 3 (56 weeks). The higher discontinuation rate in Study 3 was due, in part, to patients being discontinued at Week 16 if they did not meet pre-specified efficacy criteria to ensure (for safety purposes) that only patients receiving an efficacy benefit received longer-term treatment [14].

3.2. Median time to first improvement

Fig. 1A shows the median time to first improvement profile for WOMAC Pain in placebo-controlled Study 1 (16-week treatment). The x-axis shows improvement thresholds (i.e., % change in pain from baseline), and the y-axis shows median time to the event (i.e., time when half
of all patients reached the corresponding level of improvement on the x-axis). For example, median time to 30% improvement (x-axis) was 2 weeks (y-axis) in both tanezumab arms and 4 weeks (y-axis) in the placebo arm. Thus, median time to first improvement of 30% in Pain was shorter with tanezumab than with placebo.

Generally, median times to first improvement in Pain in Study 1 were similar across treatment arms at thresholds of improvement ≥20%. At thresholds >20%, both tanezumab arms (2.5 and 2.5/5 mg) exhibited shorter median times to first improvement in Pain than the placebo arm. For example, median time to first improvement of 50% was 2 weeks in both tanezumab arms and 12 weeks in the placebo arm. Comparisons between tanezumab and placebo were statistically significant for all thresholds ≥20% (Supplementary Table 1A, left panel). Median times to first improvement in Pain >50% could not be estimated for placebo, since not enough (i.e., less than half) patients in the placebo arm achieved these levels of improvement. In contrast, median times to first improvement in Pain were still estimable at higher thresholds (up to 65–70%) in the tanezumab arms. Overall, both tanezumab dose arms exhibited similar median times to first improvement in Pain, particularly at thresholds ≤50% (it should be noted that all patients received a 2.5 mg dose until Week 8 in this study). The median time to first improvement profile for WOMAC Physical Function (Fig. 1B) and Stiffness (Fig. 1C) in Study 1 were similar to the Pain domain. All comparisons between tanezumab and placebo were significant for Physical Function (Supplementary Table 1A, middle panel) and Stiffness (Supplementary Table 1A, right panel). Similarity among the three WOMAC domains is evident in Supplementary Fig. 1A, which shows the median time to first improvement profile for each domain for the tanezumab 2.5/5 mg group. Times to first improvement of 30% and 50% were 2 and 4 weeks, respectively, for all WOMAC domains.

Findings from placebo-controlled Study 2 (24-week treatment) generally followed the same pattern as Study 1. Across all WOMAC domains, median times to first improvement were typically shorter in the tanezumab arms (2.5 and 5 mg) than the placebo arm at thresholds >15% (Fig. 2A–C), median times to first improvement were not estimable for the placebo arm at higher thresholds (>55%), and both tanezumab doses produced generally similar median time-to-improvement patterns. Across all domains, all comparisons between tanezumab and placebo were statistically significant (Supplementary Table 1B). In contrast to Study 1, we observed that median times to first improvements were somewhat shorter in the Stiffness domain than in the Pain and Physical Function domains, particularly in the tanezumab 5 mg arm in Study 2. In the tanezumab 5 mg arm, for example, median time to first improvement of 50% was 4 weeks for Stiffness but 12 weeks for Pain and Physical Function (Supplementary Fig. 1B).

In active-controlled Study 3 (56-week treatment), there were no significant differences between the NSAID and tanezumab (2.5 and 5 mg) arms in terms of median times to first improvement of any WOMAC domain at any threshold of improvement (Fig. 3A–C; Supplementary Table 1C). Due to the longer treatment period in this study (56 weeks) relative to the placebo-controlled studies (16–24 weeks), and the presence of an active comparator, median times to improvement were estimable at higher thresholds (75–80%) in all treatment arms. There was little difference in the pattern of response among the different WOMAC domains in the tanezumab 5 mg arm (Supplementary Fig. 1C); median
times to first improvement of 30% and 50% in response to tanezumab 5 mg were 4 weeks and 16 weeks, respectively, for all domains.

### 3.3. Median time to sustained improvement

Across all WOMAC domains and treatment arms in Study 1, most patients experiencing an event of first improvement also experienced a sustained event. For example, 69.8%, 81.0%, and 82.0% of all patients in the placebo, tanezumab 2.5/5 mg, and tanezumab 5 mg arms, respectively, experienced an event of first improvement in Pain of 30%. Of these patients, 95.7%, 93.0%, and 92.7%, respectively, experienced a sustained event (Table 2). Across all domains, median times to sustained improvement were shorter in the tanezumab arms than in the placebo arm, with clear separation between the arms at thresholds >15% (Fig. 1D-F). Comparisons between tanezumab and placebo were significant for all but one threshold (the 5% threshold for the Pain domain; Supplementary Table 2A). Although minor differences were noted, both doses of tanezumab (2.5 and 2.5/5 mg) exhibited generally similar median times to sustained event across all domains. Likewise, there was little difference in the pattern of response across the different WOMAC domains in the tanezumab 2.5/5 mg arm, particularly at thresholds <50% (Supplementary Fig. 2A).

Findings from Study 2 were similar to those from Study 1. Most patients with an event of first improvement also experienced sustained improvement (Table 3). Both tanezumab groups (2.5 and 5 mg) exhibited shorter median times to sustained improvement in all domains than the placebo group (Fig. 2D-F). All comparisons between tanezumab and placebo were significant (Supplementary Table 2B), and there was little difference in the pattern of response between tanezumab doses. Minor differences were observed between the different WOMAC domains, but overall the pattern of response in the tanezumab 5 mg arm was generally similar across the different domains (Supplementary Fig. 2B).

As in Studies 1 and 2, most patients who experienced an event of first improvement also experienced sustained improvement in Study 3 (Table 4). Median times to sustained improvement in Study 3 were similar between the NSAID and tanezumab arms across all WOMAC domains and thresholds of improvement (Fig. 3D-F). The only comparisons that were significantly different between NSAID and tanezumab were the 35% and 40% thresholds for the Pain domain (Supplementary Table 2C). There was little difference in the pattern of response among the WOMAC domains in the tanezumab 5 mg arm (Supplementary Fig. 2C).

Generally, median times to sustained improvement in Pain were similar to median times to first improvement at lower thresholds (<25–40% across studies) in each study (Supplementary Fig. 3). In Study 1, for example, median time to first improvement and median time to sustained improvement in Pain of 30% were 2 weeks. At higher thresholds, median times to first improvement were shorter than median times to sustained improvement. In Study 1, for example, median time to first improvement in Pain of 50% was 4 weeks, but median time to sustained improvement of 50% was 12 weeks. At thresholds >50%, median time to sustained improvement was not always estimable, since not enough patients (less than half) achieved this level of sustained improvement. In Study 1, for example, 50.2% of patients in the tanezumab 2.5/5 mg arm experienced an event of first improvement in Pain of 70% (Table 2), with a median time to improvement of 16 weeks (Supplementary Fig. 3). However, median time to sustained improvement was not estimable for the tanezumab 2.5/5 mg arm, since less than half of all patients in this arm (42.1%; Table 2) achieved a sustained improvement of 70%. Similar findings were observed for the other WOMAC domains in each study (data not shown).

### 4. Discussion

This study utilized a novel approach to characterize onset of treatment effect with tanezumab in patients with OA by examining the relationship between pre-specified thresholds of symptomatic improvement and the estimated median times to achieve those thresholds across three separate clinical trials. We found that median time to first improvement in all WOMAC domains was consistently and notably shorter among tanezumab-treated patients than in patients treated with placebo for thresholds of improvements >15–20%. Across all WOMAC domains, no noteworthy differences in median time to first improvement were observed between tanezumab- and NSAID-treated patients for all domains.
thresholds of improvement. With the exception of the Stiffness domain in Study 2 (which exhibited somewhat shorter median times to improvement across most thresholds relative to Pain and Physical Function), there was little difference in the pattern of response (across all thresholds) for the different WOMAC domains. Likewise, there was little difference in the pattern of response for different doses of tanezumab in the individual studies (i.e., tanezumab 2.5 mg was similar to tanezumab 2.5/5 mg and tanezumab 5 mg), particularly at thresholds of improvement <40–50%. Across all domains, most patients who experienced an event of first improvement also experienced sustained improvement. Events of sustained improvement typically occurred relatively quickly after (or simultaneously to) events of first improvement at lower thresholds (up to 40–45% across studies). At higher thresholds, median time to a sustained improvement event was typically longer than that of first improvement events, suggesting that patient response may fluctuate more before stabilizing at these thresholds. Generally, median times to first (but not sustained) improvement of >50% were not always estimable across all studies and WOMAC domains, suggesting that some patients may hit a high (but not necessarily sustained on average) point of symptomatic improvement.

Despite the similarities observed across studies, there was some variability in estimations of median time to improvement among tanezumab-treated patients, particularly at higher thresholds of improvement. For example, whereas median time to first improvement of 30% (a level considered moderate and clinically meaningful among patients with chronic pain [25]) ranged from 2 to 4 weeks across all WOMAC domains and studies, variation in median time to first improvement of 50% (considered substantial improvement [25]) was more marked. Median time to first improvement of 50% was 4 weeks for all domains in Study 1, 12 weeks for Pain and Physical Function (4–8 weeks for Stiffness) in Study 2, and 16 weeks for all domains in Study 3. Across all WOMAC domains, median time to sustained improvement of 30% was 2–4 weeks in Study 1, 4 weeks in Study 2, and 4–8 weeks in Study 3. Median time to sustained improvement of 50% was 8–12 weeks in Study 1, 12–24 weeks in Study 2, and 16 weeks in Study 3. Such variability is expected across studies and may be explained by differences in study design and/or patient populations. For example, the proportion of white patients was higher in Study 2 (87%) than in the other studies (70–72%). There were also differences in treatment duration (Study 1 = 16 weeks, Study 2 = 24 weeks, Study 3 = 56 weeks), geography (Study 1 = North America, Study 2 = Europe and Japan, Study 3 = North America, South America, Europe, and Asia-Pacific), and eligibility criteria (Studies 1 and 2 required inadequate response or intolerance to acetaminophen, NSAIDs, trandolapril and opioids, whereas Study 3 had similar requirements for acetaminophen, trandolapril, and opioids but required patients to be on a sustained, tolerable dose of NSAIDs at screening).

Tanezumab's onset of effect in Studies 1 and 2 has been examined previously, based on the first time point where mean change (from baseline) in WOMAC score was significantly greater in the tanezumab arms than in the placebo arm. Onset of effect for all tanezumab arms was estimated at 2 weeks for all WOMAC domains assessed (Study 1: Pain and Physical Function; Study 2: Pain, Physical Function, and Stiffness) [12, 26]. Improvements over placebo were evident at most (Study 1) or all (Study 2) time points up to the end of the treatment, but there were no formal criteria for defining sustained improvement as in the current analysis [12, 26]. Our findings from Study 1 are in general agreement with these previous results in that we also observed onset of improvement in all WOMAC domains in the tanezumab arms within a few weeks of initiating treatment; median time to first improvement was 2 weeks for the 30% threshold and 4 weeks for the 50% threshold. We also observed early onset of effect in Study 2; median time to first improvement was 2–4 weeks for the 30% threshold and 4 weeks for the 50% threshold. We also observed early onset of effect in Study 2; median time to first improvement was 2–4 weeks for the 30% threshold. Higher thresholds had longer median times to first improvement in Study 2; 4–12 weeks for the 50% threshold. Our median time to first improvement estimates (based on WOMAC domains) also agree with a previous pooled analysis of Studies 1 and 2, which assessed median time to onset of effect of tanezumab based on changes in daily pain score (on a numeric rating scale from 0 = no pain to 10 = worst possible pain). Median time to achieve 30% and 50% improvement in daily pain was estimated at 3–4 weeks and 11 weeks, respectively [27].

Previous studies have estimated a relatively quick time to onset of effect for NSAIDs in patients with OA (2–5 days) [28, 29]. For a variety of reasons, it is difficult to directly compare time-to-onset estimates from previous studies with estimates for NSAIDs and for tanezumab from our study. First, our study lacked data prior to Week 2, so we are unable to assess onset of effect at earlier timepoints. Second, previous studies typically base response on statistical difference from placebo using group mean treatment effects, whereas our analysis based response on 50% of

---

**Table 3**

Proportion of patients experiencing first and sustained improvement events in the WOMAC domains during study 2.

| N         | Pain                          | Physical Function                  | Stiffness                        |
|-----------|-------------------------------|-----------------------------------|----------------------------------|
|           | First n (%) | Sustained n (%) | First to Sustained (%) | First n (%) | Sustained n (%) | First to Sustained (%) | First n (%) | Sustained n (%) | First to Sustained (%) |
|-----------|--------------|-----------------|----------------------------|--------------|-----------------|------------------------|--------------|-----------------|------------------------|
| PBO       | 282          | 242 (85.8)      | 222 (78.7)                | 91.7         | 245 (86.9)      | 220 (78.0)            | 89.8         | 247 (87.6)      | 220 (78.0)            |
|           | 283          | 268 (94.7)      | 258 (91.2)                | 96.3         | 268 (94.7)      | 253 (89.4)            | 94.4         | 263 (92.9)      | 250 (88.3)            |
|           | 284          | 271 (95.4)      | 252 (88.7)                | 93.0         | 262 (92.3)      | 252 (88.7)            | 96.2         | 269 (94.7)      | 250 (88.0)            |
| 30%       | PBO          | 282             | 207 (73.4)                | 186 (66.0)   | 89.9            | 196 (69.5)            | 170 (60.3)   | 86.7            | 209 (74.1)            |
|           | 283          | 245 (86.6)      | 220 (77.7)                | 89.8         | 238 (81.4)      | 214 (75.6)            | 89.9         | 250 (88.3)      | 219 (77.4)            |
|           | 284          | 238 (83.8)      | 215 (75.7)                | 90.3         | 231 (81.3)      | 217 (76.4)            | 93.9         | 241 (84.9)      | 224 (78.9)            |
| 50%       | PBO          | 282             | 152 (53.9)                | 119 (42.2)   | 78.3            | 137 (48.6)            | 107 (37.9)   | 78.1            | 160 (56.7)            |
|           | 283          | 192 (67.8)      | 154 (54.4)                | 80.2         | 171 (60.4)      | 142 (50.2)            | 83.0         | 207 (73.1)      | 160 (56.5)            |
|           | 284          | 194 (68.3)      | 171 (60.2)                | 88.1         | 181 (63.7)      | 153 (53.9)            | 84.5         | 211 (74.3)      | 172 (60.6)            |
| 70%       | PBO          | 282             | 83 (29.4)                 | 63 (22.3)    | 75.9            | 67 (23.8)             | 48 (17.0)    | 71.6            | 95 (33.7)             |
|           | 283          | 107 (37.8)      | 76 (26.9)                 | 71.0         | 92 (32.5)       | 67 (23.7)             | 72.8         | 135 (47.7)      | 91 (32.2)             |
|           | 284          | 113 (39.8)      | 80 (28.2)                 | 70.8         | 97 (34.2)       | 66 (23.2)             | 68.0         | 152 (53.5)      | 110 (38.7)            |

PBO = placebo, TNZ = tanezumab.

* The 15%, 30%, 50%, and 70% thresholds were selected as representative, clinically meaningful thresholds of improvement.

b First to sustained shows, among patients experiencing a first event, the proportion who experienced a sustained event (calculated as sustained n divided by first n).
patients achieving a pre-set threshold of improvement. Third, previous studies typically assess time to first improvement without any assessment of whether sustained improvement occurs. Finally, in contrast to previous studies, the current analysis was carried out in a different population (i.e., patients with moderate to severe OA and a history of inadequate response or intolerance to standard analgesics) that may exhibit different time-to-onset profiles than less severe patients or patients that respond adequately to other OA treatments.

Conclusions from our analyses are somewhat limited by study design. Post-baseline WOMAC assessments were completed intermittently during each study. This affects the precision of median time-to-improvement estimates, particularly at later time points in Studies 2 (after Week 16) and 3 (after Week 8) when the time between assessments was 8 weeks. As mentioned earlier, the first post-baseline WOMAC assessment was not until Week 2. However, statistically significant (compared to placebo) improvements in pain (based on daily pain scores using a numeric rating scale from 0 = worst possible pain) occur within days of initiating tanezumab treatment [12]. Therefore, there may be notable WOMAC findings at early time points (differences between the NSAID and tanezumab groups, for example) that could not be captured in our analysis. The treatment duration of the placebo-controlled studies (16–24 weeks) limits our estimation of time to sustained improvement, since it is not known if patients would have maintained the specified levels of improvement for longer time periods. In Study 1, for example, median time to sustained improvement of 50% in Pain was estimated at 8–12 weeks. Our definition of sustained improvement (improvement that was sustained through the end of treatment), however, only required patients to maintain this level of improvement for an additional 4 weeks, since the treatment period ended at Week 16. Findings around active-controlled Study 3 are limited at later time points, since nearly 60% of patients discontinued the 56-week treatment period at some point. This is particularly relevant after Week 16 when patients had to meet specific efficacy criteria to remain in the study [14].

Pre-specified efficacy discontinuities at Week 16 affected all treatment arms equally (21–22% of patients) but could have confounded potential treatment differences between the NSAIDs and tanezumab arms, since the study was enriched for patients who responded well to treatment. Finally, the population included in our study consisted of patients with moderate to severe OA and a history of inadequate response or intolerance to analgesics commonly used to manage OA. Results may not be generalizable to patients with milder symptoms or who respond adequately to other treatments. However, the study population can also be viewed as a strength of our analysis, since we were able to demonstrate clear differences in median time to symptomatic improvement (pain, function, and stiffness) between tanezumab and placebo in this difficult-to-treat patient population. Another strength is the large sample size of our analysis, which included 514 placebo-treated patients, 3031 tanezumab-treated patients, and 996 patients treated with NSAIDs.

Overall, our novel approach utilizing nonparametric time-to-event analyses demonstrates that tanezumab-treated patients generally have a shorter time-to-onset of effect and may reach higher thresholds of improvement than placebo-treated patients. Patients with moderate to severe OA and a history of inadequate response or intolerance to standard analgesics achieved first improvement of symptoms (pain, physical function, and stiffness) of 30% within 2–4 weeks of initiating treatment with tanezumab and sustained improvement of 30% within 2–8 weeks of initiating treatment. Median time to improvement of 50% was more variable, with first and sustained improvement evident within 4–16 and 8–24 weeks, respectively. We found no notable differences in the median time-to-improvement profile of tanezumab relative to NSAIDs, suggesting that the two agents have a similar impact in this particular study population (i.e., patients who were receiving sustained NSAIDs prior to trial enrollment) after 2 weeks of treatment (the first time point assessed). An understanding of median time-to-improvement profiles, such as those presented here for tanezumab and NSAIDs, can help physicians and patients set realistic OA treatment expectations, which could improve adherence and, potentially, treatment outcomes.

**Ethics approval**

The studies included in this analysis were conducted in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration, and all patients provided informed consent.
Role of the funding source

The study was sponsored by Pfizer and Eli Lilly and Company. Pfizer, in collaboration with Eli Lilly and Company, is the manufacturer of tanezumab. Pfizer Inc and Eli Lilly and Company contributed to the study design; Pfizer contributed to the management and collection of data. In their role as authors, employees of Pfizer and Eli Lilly were involved in the interpretation of data, preparation, review, and approval of the manuscript and the decision to submit for publication, along with their co-authors. The study sponsors approved the manuscript from an intellectual property perspective but had no right to veto the publication.

Author contribution

All authors contributed to (1) the conception and design of the study, or acquisition of data, or analysis and interpretation of data, (2) drafting the article or revising it critically for important intellectual content, and (3) final approval of the version to be submitted.

Competing interests

DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, and Novartis. TJS serves as a consultant to Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, P

Author contribution

All authors contributed to (1) the conception and design of the study, or acquisition of data, or analysis and interpretation of data, (2) drafting the article or revising it critically for important intellectual content, and (3) final approval of the version to be submitted.

Competing interests

DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, and Novartis. TJS serves as a consultant to Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunenthal, Techfields, Galapagos, and Unity; has served on Data Safety Monitoring Boards for IQVIA and Astra-Zeneca; and has received grants or contract support from IQVIA, Vertx, GlaxoSmithKline, Collegium, Acadia, and Grunenthal. DJH serves as a consultant to Tissuegene, TLC Bio, Pfizer, Eli Lilly and Company, Vertx, GlaxoSmithKline, Collegium, Acadia, Grunential...
[27] T.J. Schnitzer, F. Berenbaum, I. Davignon, R. Yang, L. Viktrup, C.R. West, et al., Evaluating analgesic response to subcutaneous tanezumab in patients with inadequate treatment response to other analgesics based on daily e-pain diaries: a pooled analysis of 2 randomized, placebo-controlled studies [abstract], Arthritis Rheumatol. 72 (2020) 3317–3319. https://acrabstracts.org/wp-content/uploads/2020/11/2020ACRAbstract-Supplement.pdf.

[28] R.J. Holt, J.G. Fort, A.Y. Grahn, J.D. Kent, A.E. Bello, Onset and durability of pain relief in knee osteoarthritis: pooled results from two placebo trials of naproxen/esomeprazole combination and celecoxib, Phys Sportsmed 43 (2015) 200–212, https://doi.org/10.1080/00913847.2015.1074852.

[29] W.P. Battisti, N.P. Katz, A.L. Weaver, A.K. Matsumoto, A.J. Kivitz, A.B. Polis, et al., Pain management in osteoarthritis: a focus on onset of efficacy—a comparison of rofecoxib, celecoxib, acetaminophen, and nabumetone across four clinical trials, J. Pain 5 (2004) 511–520, https://doi.org/10.1016/j.jpain.2004.09.004.