HHS Public Access

Author manuscript

Osteoarthritis Cartilage. Author manuscript; available in PMC 2017 June 01.

Published in final edited form as:

Osteoarthritis Cartilage. 2016 June ; 24(6): 962–972. doi:10.1016/j.joca.2016.01.135.

Comparative pain reduction of oral non-steroidal anti-inflammatory drugs and opioids for knee osteoarthritis: systematic analytic review

Savannah R. Smith, BA[1,*], **Bhushan R. Deshpande, BS**[1,*], **Jamie E. Collins, PhD**[1,2], **Jeffrey N. Katz, MD, MSc**[1,2,3,4], and **Elena Losina, PhD**[1,2,3,5]

Savannah R. Smith: ssmith94@partners.org; Bhushan R. Deshpande: bdeshpande@partners.org; Jamie E. Collins: jcollins13@partners.org; Jeffrey N. Katz: jnkatz@partners.org; Elena Losina: elosina@partners.org

[1]Orthopaedic and Arthritis Center for Outcomes Research, Department of Orthopaedic Surgery, Brigham and Women's Hospital, Boston University School of Public Health – all Boston, MA

[2]Harvard Medical School, Immunology and Allergy, Brigham and Women's Hospital, Boston University School of Public Health – all Boston, MA

[3]Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston University School of Public Health – all Boston, MA

[4]Departments of Epidemiology and Environmental Health, Harvard T. H. Chan School of Public Health, Boston University School of Public Health – all Boston, MA

[5]Department of Biostatistics, Boston University School of Public Health – all Boston, MA

Abstract

Objective—Summarize the comparative effectiveness of oral non-steroidal anti-inflammatory drugs (NSAIDs) and opioids in reducing knee osteoarthritis (OA) pain.

Correspondence: Elena Losina, PhD, Orthopaedic and Arthritis Center for Outcomes Research, Department of Orthopedic Surgery, Brigham and Women's Hospital, 75 Francis St, BC 4-016, Boston, MA 02115, elosina@partners.org, Phone: (617) 732-5338, Fax: (617) 525-7900.

*Savannah R. Smith, BA, and Bhushan R. Deshpande, BS, contributed equally to the preparation of this article.

Author Contributions: Drs. Losina and Collins had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Conception and design: Smith, Deshpande, Losina
Collection and assembly of data: Smith, Deshpande
Analysis and interpretation of the data: Smith, Deshpande, Collins, Katz, Losina
Statistical expertise: Collins, Losina
Drafting of the article: Smith, Deshpande, Collins
Critical revision of the article for important intellectual content: Smith, Deshpande, Collins, Katz, Losina
Final approval of the article: Smith, Deshpande, Collins, Katz, Losina
Obtaining of funding: Katz, Losina

Previous presentation: A portion of the study was presented at the May 2015 World Congress of the Osteoarthritis Research Society International.

Competing interests: Dr. Katz is the President-Elect of the Osteoarthritis Research Society International. Drs. Katz and Losina are Deputy Editors for Biostatistics and Methodology for the *Journal of Bone and Joint Surgery*. All other authors have no conflicts of interest to report.

Methods—Two reviewers independently screened reports of randomized controlled trials, published in English between 1982 and 2015, evaluating oral NSAIDs or opioids for knee OA. Included studies were at least eight weeks duration, conducted in Western Europe, the Americas, New Zealand, or Australia, and reported baseline and follow-up pain using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) Pain subscale (0–100, 100-worst). Effectiveness was evaluated as reduction in pain, accounting for study dropout and heterogeneity.

Results—27 treatment arms (9 celecoxib, 4 non-selective NSAIDs [diclofenac, naproxen, piroxicam], 11 less potent opioids [tramadol], and 3 potent opioids [hydromorphone, oxycodone]) from 17 studies were included. NSAID and opioid studies reported similar baseline demographics and efficacy withdrawal rates; NSAID studies reported lower baseline pain and toxicity withdrawal rates. Accounting for efficacy-related withdrawals, all drug classes were associated with similar pain reductions (NSAIDs: −18; less potent opioids: −18; potent opioids: −19). Meta-regression did not reveal differential effectiveness by drug class but found that study cohorts with a higher proportion of male subjects and worse mean baseline pain had greater pain reduction. Similarly, results of the network meta-analysis did not find a significant difference in WOMAC Pain reduction for the three analgesic classes.

Conclusion—NSAIDs and opioids offer similar pain relief in OA patients. These data could help clinicians and patients discuss likely benefits of alternative analgesics.

Keywords

Knee osteoarthritis; NSAIDs; Opioids; WOMAC Pain subscale; Meta-analysis; Meta-regression; Network meta-analysis

Introduction

Knee osteoarthritis (OA) affects millions of American adults and is characterized by substantial pain, joint stiffness, and functional limitations.[1] Although over half of all knee OA patients eventually undergo total knee replacement, nearly all will require at least some amount of long-term pain control.[2]

Standard treatment begins with non-pharmacologic approaches to symptom relief and functional restoration, including weight reduction, orthotic devices, exercise, and physical therapy. Because these treatments often provide limited pain relief, pharmacologic analgesics are frequently also employed. Many professional societies suggest the use of non-steroidal anti-inflammatory drugs (NSAIDs) or tramadol, a lower potency opioid, for primary pharmacologic management of knee OA. Recommendations on the use of more potent opioids remain conflicted for this population.[3–6]

Both NSAIDs and opioids are associated with a wide variety of adverse effects, and there are no long-term trials of knee OA patients comparing their efficacies. Given the limited comparative evaluation of oral NSAIDs and opioids in the knee OA population as well as the importance of understanding the long-term effectiveness of the each of the classes of drugs through decision modeling and formal comparative-effectiveness analyses, we employed a systematic analytic review to evaluate both classes of analgesics in reducing pain among persons with OA.

Methods

We conducted our analysis according to the principles of PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.[7] Our analysis was not preregistered.

Identification of studies

We conducted a search of all articles available in PubMed, Web of Science – Science Citation Index Expanded, EMBASE, and the Cochrane Central Register of Controlled Trials. Our search identified all articles including *osteoarthritis* and any of the following terms in the title: *non-steroidal anti-inflammatory drug*(*s*), *NSAID*(*s*), *ibuprofen*, *celecoxib, diclofenac*, *naproxen*, *meloxicam*, *nabumetone*, *etodolac*, *indomethacin*, *piroxicam*, *sulindac*, *salsalate*, *flurbiprofen*, *ketoprofen*, *oxycodone, hydrocodone*, *hydromorphone*, *fentanyl*, *methadone*, *morphine*, *tramadol*, or *codeine*. These search terms reflect NSAIDs and opioids commonly prescribed to US Medicare beneficiaries with knee OA as of 2009,[8] excluding the since-withdrawn propoxyphene.[9] Two reviewers (SRS and BRD) independently screened the abstract of each article to determine whether it was a randomized controlled trial (RCT) conducted in humans and published between January 1, 2000 and March 6, 2015.

Inclusion and exclusion criteria

We included clinical trials of predominantly knee OA patients of at least eight weeks duration that evaluated efficacy of oral analgesics using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) Pain subscale;[10] study arms that combined patients with OA and patients with other forms of arthritis were excluded. We included only reports of RCTs, as other study designs do not restrict concurrent treatment utilization and would not provide measures of pain severity pre- and post-treatment initiation. As more invasive placebos have been associated with greater effectiveness,[11,12] we excluded studies employing combination therapies of oral analgesics and non-oral placebos to provide a more homogeneous basis for our direct and indirect comparisons. Studies not reporting group mean and standard deviation values for baseline pain and either change from baseline or follow-up pain were excluded for insufficient data. When possible, we calculated standard deviations from reported standard errors or confidence intervals. Additional exclusion criteria eliminated studies that were not published in English or were primarily conducted outside of developed countries (defined as Western Europe, the Americas, New Zealand, or Australia). For studies that reported multiple follow-up time points, we selected the one nearest twelve weeks.

Data abstraction and quality assessment

From the reports meeting all inclusion criteria, we obtained the following data: identity and dose of drugs evaluated; funding source; geographic location of the study; sample size; discontinuations due to loss of efficacy, adverse events (clinical adverse event, laboratory adverse event, and fatality), and other reasons; cohort characteristics (age, gender, height, weight, body mass index, time since OA diagnosis, race/ethnicity, and primary joint affected [knee vs. hip]); study duration; baseline WOMAC Pain (mean, standard deviation); and either change in WOMAC Pain (mean, standard deviation), follow-up WOMAC Pain (mean,

standard deviation) or both. If the range or directionality of the scale was ambiguous, we contacted the authors for clarification. Except where noted, all abstracted data were obtained from the intention-to-treat (ITT) analysis population. Included articles were evaluated for quality using the Jadad assessment tool, a 5-point scoring system assessing reports of RCTs based on appropriate methods of randomization, blinding, and withdrawal reporting.[13]

The two reviewers independently completed all screening, data extraction, and quality assessment. Cases of disagreement were discussed and resolved by the two reviewers, consulting other authors if necessary. In a sensitivity analysis, we excluded study arms evaluating 100mg and 400mg tramadol, which are not representative of contemporary clinical practice.

Statistical analysis

We converted pain data to a 0–100 (100 worst) scale by arithmetic transformation[14] and evaluated cohort differences using the t-approximation of the Wilcoxon rank sum test.

For studies providing only baseline and final pain scores, we calculated mean change by subtracting final pain from baseline pain. To calculate the standard deviation of change we first calculated both the correlation between baseline pain and change in pain as well as the correlation between baseline pain and follow-up pain for the studies that reported all three time points. We calculated the standard deviation of change for those studies that did not explicitly report it, using properties of variances (the variance of the sum of two distributions is the sum of the distributions' variances plus twice their covariance) and assuming that the correlations derived among studies that reported all three time points would also apply for those that reported only two time points.[15] Finally, we modified the mean change in pain accounting for withdrawals due to insufficient efficacy by assuming, conservatively, that these subjects would report no change from baseline. While we abstracted data from ITT analyses, which employ methods to handle missing data, the specific methods used were heterogeneous, including strategies such as last observation carried forward, baseline observation carried forward, and imputation utilizing dropout reason. Thus, we modified mean change in pain to account for inefficacy withdrawals to produce a more conservative estimate. In sensitivity analyses, we used the unadjusted mean change in pain as reported in the literature. The standard error of change was calculated by dividing the standard deviation by the square root of the sample size, using the intention to treat population when reported and the number randomized otherwise. Mean changes were combined into a weighted average, weighted by the precision (the reciprocal of the variance) of the each estimate. Separate analyses were performed for three analgesic classes: NSAIDs, less potent opioids, and potent opioids.

We used funnel plots and Egger's linear regression test to investigate publication bias. We chose Egger's test over the rank correlation test because the rank test has been shown to have low power when the number of studies is small.[16,17] When publication bias was suspected, we used the trim and fill method as a sensitivity analysis.[18] The trim and fill is a non-parametric method to correct for publication bias. It uses rank-based augmentation techniques to impute potential missing studies in order to make the funnel plot symmetric. Outcomes are re-estimated on the augmented data. To determine if the results were robust to

assumptions of the meta-analysis, we performed heterogeneity analyses and report the H and I^2 statistics for each analysis.[19] The contribution of each study to the overall heterogeneity was assessed by the Q-term and influence. The Q-term is the contribution of the study to Cochrane's Q statistic, and the influence is computer by comparing the overall pooled estimate with and without the study included. We used a random effects analysis using restricted maximum likelihood to calculate a final combined estimate of change in pain in order to account for heterogeneity. Finally, we used meta-regression to determine factors systematically associated with greater change in pain.[20] We included a 3-level analgesic class variable (NSAIDs vs. less potent opioids vs. potent opioids) as the primary independent variable of interest and adjusted for mean baseline pain, percent of the cohort with knee OA (vs. hip), percent of the cohort that was female, study year, and country (exclusively US-based vs. all other).

We conducted a secondary network meta-analysis (NMA), using the NMA framework to evaluate both direct and indirect comparisons between NSAIDs and opioids. We used a random-effects model with Gaussian quadrature to fit the model.[21,22]

All statistical analyses were conducted using SAS 9.4 (SAS Institute Inc., Cary, NC).

Results

Study selection

Figure 1 summarizes the article selection process. Our initial search identified 1,688 unique articles: 1,535 evaluating NSAIDs and 153 evaluating opioids. Upon screening the abstracts of the 940 articles published post-1999, we identified 247 articles for full text review (209 and 38 for NSAIDs and opioids, respectively). Of those, 24% (60/247) were excluded for not utilizing the WOMAC Pain subscale. Additionally, 10% (25/247) measured pain via the WOMAC Pain subscale, but did not report sufficient data to be included in our analysis. Six articles (2%) were excluded for their use of non-oral placebo. We identified 17 studies meeting all outlined inclusion criteria: 11 examining just NSAIDs (celecoxib, diclofenac, naproxen, piroxicam), 3 examining just less potent opioids (tramadol, tramadol/acetaminophen), 1 examining both NSAIDs and less potent opioids (celecoxib, tramadol), and 2 examining potent opioids (hydromorphone, oxycodone). This resulted in 27 active treatment arms to be included in our analysis (celecoxib [9], diclofenac [1], naproxen [2], piroxicam [1], tramadol [10], tramadol/acetaminophen [1], hydromorphone [2], and oxycodone [1]).

Table 1 describes the included studies with selected abstracted data. Trial duration ranged from 8 to 52 weeks; median duration was 13 weeks for NSAID treatment arms and 12 weeks for opioid treatment arms ($p<0.01$). The size of the treatment arms varied from 25 to 481 (median 236) persons for NSAID arms and from 60 to 202 (median 176) persons for opioid arms ($p<0.01$). Several baseline patient demographics did not vary substantially from a clinical perspective for NSAID and opioid studies, with a median age of 62 years for the NSAID arms compared to 60 years for the opioid arms. Mean baseline WOMAC Pain was somewhat lower for NSAID arms (52 points) than opioid arms (60 points, $p=0.04$). Treatment arms evaluating NSAIDs reported shorter median time-since-diagnosis (5 years

and 8 years, respectively; p=0.03), as well as a lower median BMI (31.0 kg/m^2 and 32.4 kg/m^2 respectively; p=0.07) and proportion of subjects withdrawing due to toxicity compared to opioid studies (7% vs. 24%, respectively, p<0.01). NSAIDs and opioid studies presented a similar median proportion of subjects withdrawing due to insufficient efficacy (7% vs. 11%, respectively, p=0.10) or any other reason (5% vs. 8%, respectively, p=0.40).

Table 2 shows change in WOMAC Pain values modified to account for subjects withdrawing due to insufficient efficacy, by assuming these subjects would report no change from baseline.

Among the 17 included articles, 12 (71%) had a Jadad quality score of 4 or 5 (maximum score 5), and the remainder had a score of 3. All articles detailed the withdrawals and dropouts during the trials, and only two articles did not report any funding from the pharmaceutical industry.

Heterogeneity and effectiveness

NSAIDs (Celecoxib, Diclofenac, Naproxen, Piroxicam)—NSAIDs studies exhibited a large amount of heterogeneity (I^2=0.95, H=4.58). The estimate from Hochberg et al.[23] was the most influential (Influence=9.35) and contributed substantial weight to the heterogeneity score (Q-term=150). We investigated whether dropping the Hochberg study would reduce heterogeneity, as it had the largest pain decrement among NSAID studies (adjusted WOMAC Pain change = −35.6) and the highest baseline pain (74.1). Heterogeneity remained high after excluding Hochberg et al.[23] (I^2=0.88, H=2.90); thus, we included this study and further investigated sources of heterogeneity in meta-regression.

The high heterogeneity suggested that a fixed effects approach was inappropriate, and we therefore used a random effects analysis. The random effects model, accounting for between-observation and between-study variability, produced a combined estimate of −18 (SE 1.9) (Figure 3a). In a sensitivity analysis using reported unadjusted mean change in pain, we estimated a pain decrement of −20 (SE 2.1).

Less Potent Opioids (Tramadol, Tramadol/Acetaminophen)—The analysis of heterogeneity suggested moderate to high inconsistency and heterogeneity (I^2=0.71, H=1.85). The most influential study arms were the 100mg tramadol dose in DeLemos et al.,[24] the tramadol/acetaminophen treatment group Emkey et al.,[25] and the 300mg tramadol dose in Fishman et al.[26] (Influence=1.12, 0.75, and 0.59). These studies contributed substantial weight to the heterogeneity score (Q-term=10.1, 4.8 and 10.5). We investigated whether excluding Emkey et al.,[25] the only study not evaluating tramadol exclusively, would reduce heterogeneity and continued to find moderate to high inconsistency and heterogeneity (I^2=0.69, H=1.8). Using a random effects analysis we found a combined estimate of effectiveness of −18 (SE 1.0) for less potent opioids (Figure 3b). In sensitivity analysis using reported unadjusted mean change in pain, we estimated a pain decrement of −21 (SE 1.0).

In another sensitivity analysis, we excluded the 100mg and 400mg doses of tramadol, as they represent doses not regularly used in clinical practice, along with the combined

tramadol/acetaminophen regimen, which produced an overall pain reduction estimate of −19 (SE 1.3).

The funnel plot for less potent opioids exhibited asymmetry, suggesting that there may be missing studies which would have reported less change (Figure 2); Egger's test was not statically significant (p=0.09).The trim and fill method was used to impute hypothetical missing publications.[18] A funnel plot with imputed trim and fill values is shown in Figure 4. While the peak remained uncentered, the plot was more symmetric and Egger's test was no longer statistically significant (p=0.57). After the trim and fill imputation, the combined estimate for change in pain from baseline decreased to −17 (SE 1.0).

Potent Opioids (Oxycodone, Hydromorphone)—With only three studies, heterogeneity was difficult to assess for potent opioids. The corresponding statistics indicated low to mild inconsistency and heterogeneity (I^2=0, H=0.6), though these measures may be inflated due to the limited number of studies.[19] The estimate of the change in pain obtained from both the random and fixed effects models for potent opioids was −19 (SE 1.3) (Figure 3c). In a sensitivity analysis using reported unadjusted mean change in pain, we estimated a pain decrement of −20 (SE 1.3).

Meta-Regression

Results of the meta-regression analysis did not suggest clinically important or statistically significant difference among the drug classes under consideration (NSAIDs, less potent opioids, potent opioids, p=0.22). We found that worse mean WOMAC Pain score was significantly associated with greater amount of change in pain score (p<0.001); specifically, a 10 point higher pain score at baseline was associated with an additional 5 point decrement in WOMAC Pain score at the end of the study. Greater proportion of patients with knee (as opposed to hip) OA was associated with a greater change in WOMAC Pain, after adjusting for baseline pain (p<0.01); for example, an increase in the proportion of knee OA patients by 10% resulted in an additional 2 point decrement in WOMAC Pain.

Secondary Analysis: Network Meta-Analysis

Direct and indirect treatment comparisons are shown in Figure 5. The mean treatment effect across the 9 comparisons of placebo and NSAIDs was −8 (range −4 to –15), as compared to −6 (range 1 to –11) across the 11 comparisons of placebo and less potent opioids. The only direct comparison of placebo and potent opioids had a treatment effect of −1 (reduction of 17 points in placebo compared to 18 in potent opioids). One study directly compared less potent opioids and NSAIDs; the mean decrement in WOMAC Pain for the NSAID arm was 22, compared to 12, 15, and 21 in the Tramadol 100mg, 200mg, and 300mg arms, respectively. The network meta-analysis suggested a trend for NSAIDs to result in larger WOMAC Pain changes than opioids; however, these differences did not reach statistical significance: NSAIDs vs. less potent opioids (Δ=−3.0, p=0.13), NSAIDs vs. potent opioids (Δ=−7.5, p=0.08), less potent vs. potent opioids (Δ=−4.4, p=0.31).

Discussion

We used meta-analytic techniques to evaluate pain reduction in persons with OA treated with NSAIDs and opioids as reported in RCTs. Our results suggest that the mean decrement in WOMAC Pain achieved by NSAIDs (−18 points), less potent opioids (−18) and potent opioids (−19) are all comparable. Opioid-treated subjects generally had higher pain; adjusting for this difference, we nonetheless observed comparable pain reduction across the three analgesic classes. Clinicians must consider differences in patient populations when discussing the pain relief one can expect from NSAIDs or opioids. As it is likely that most patients considering opioids have previously taken NSAIDs, our analyses provide a practical way of describing the extent of pain relief a patient can expect with opioids.

There exists literature summarizing the effectiveness of NSAIDs and opioids in OA management; however, this is the first to focus on analgesics commonly employed in knee OA treatment and evaluate effectiveness using WOMAC Pain, the most commonly used pain instrument in an OA population. Four previous reviews of these analgesics in the OA literature were identified: Verkleij et al.,[27] evaluating short-term effects of NSAIDs and acetaminophen; Bjordal et al.,[28] evaluating short-term effects of NSAIDs and opioids; Myers et al.,[29] evaluating longer-term effects of NSAIDs and opioids compared to duloxetine; and Bannuru et al.,[30] assessing the relative efficacy of analgesics for knee OA. Due to a differences in drugs of interest, pain measurement instrument, and primary outcome, few studies included in our analyses were also included in the aforementioned reviews (one, four, nine, and eight studies overlapped with our analysis and those of Verkleij et al.,[27] Bjordal et al.,[28] Myers et al.,[29] and Bannuru et al.[30] respectively). Verkleij et al.[27] and Bjordal et al.[28] did not restrict studies according to the instrument used to measure pain, while Myers et al.[29] limited analyses to studies reporting WOMAC composite scores, which includes subscales for function and stiffness along with a subscale for pain. Bannuru et al.[30] included all studies utilizing any measure of pain, function, or stiffness, and through network meta-analysis, derived effect sizes for each analgesic, which cannot be directly compared to the absolute WOMAC Pain reductions we present. Bjordal et al.[28] reported 10mm pain decrements for both NSAIDs and opioids over placebo on the 100mm Visual-Analog Scale over a one month horizon; however, the VAS cannot be directly compared to the WOMAC Pain subscale. The meta-regression conducted by Myers et al.[29] suggested a similar association between baseline and change from baseline in WOMAC composite score as we report for WOMAC Pain.

Both NSAIDs and opioids present non-trivial risks of significant adverse events, leading to contrasting views on their appropriate use. The American College of Rheumatology and European League Against Rheumatism conditionally recommend the use of NSAIDs or tramadol as primary analgesic agents and suggest using potent opioids only when all previous treatments have failed.[4,5] The Osteoarthritis Research Society International takes a more conservative stance, stating that the appropriateness of any opioid prescription is uncertain.[6] Though differing in their recommendations, professional societies consistently stress the paucity of long-term data on efficacy and adverse effects of many analgesics, particularly potent opioids.

We acknowledge several limitations of this analysis. Restricting the literature to studies published in English may have biased our evaluation; however, studies evaluating the effects of language restrictions in systematic reviews have not found any biases commonly associated with these restrictions.[31] As cultural, ethnic, and psychosocial factors have been suggested to be important influences on pain perception and response to pain stimuli, we limited our analyses to studies primarily performed in developed countries to limit the heterogeneity of study populations.[32,33] Although topical and oral NSAIDs appear to have similar efficacies,[34] there are few topical opioid formulations and none are commonly used for arthritis pain management.[35] Thus, we limited our analyses to oral formulations to examine medications with comparable delivery mechanisms.

We included only reports of RCTs. While observational studies and pragmatic trials can be employed evaluate analgesic effectiveness, they do not restrict concomitant utilization of additional treatments, thereby not allowing for an estimate of pain reduction attributable to the analgesic of interest. We ultimately excluded 246 articles because they were not randomized controlled trials; of those excluded, only 11 were cohort studies, none of which evaluated the outcome of interest.

Publication bias can be a significant problem for assessing the quality of the clinical trials literature, particularly when analyzing data from small cohorts.[36] The asymmetry in the funnel plots led us to suspect publication bias in these data, particularly for less potent opioids. We attempted to adjust for publication bias using the trim and fill method; however these results should be interpreted as a sensitivity analysis rather than a corrected estimate, as we cannot ensure that funnel plot asymmetry is caused exclusively by publication bias.

Comorbidities are frequently associated with poorer symptom management and thus are important factors in assessing analgesic effectiveness. Studies of analgesics, however, frequently exclude persons with clinically significant comorbidities and do not systematically present the distribution of comorbidities within the study population. Similarly, more than one-third of included studies failed to report BMI or duration of OA diagnosis. We were unable to adjust for these factors in the meta-regression, and thus, those results should be interpreted with caution.

Our analyses focused on the WOMAC Pain subscale. The WOMAC is contained within the Knee injury and Osteoarthritis Outcome Score (KOOS), which could have been incorporated as an outcome measure; however, no identified studies reported KOOS Pain instead of WOMAC Pain. Prior to the development of the WOMAC, numerous measures for pain and function among OA patients were and continue to be commonly used. We ultimately excluded a substantial proportion of otherwise eligible studies due to the use of another pain assessment measure. Expanding our analyses to include additional pain metrics would increase the number of eligible studies, potentially reducing heterogeneity and increasing the generalizability of our results. However, while various measures of OA pain are correlated,[37] there are no direct methods to transform a non-WOMAC measure into a validated score standardized with the WOMAC Pain subscale. Our analyses focused on the absolute pain decrements achieved from analgesics. We recognize that established methods such as the standard mean difference or effect size can be used to synthesize data from studies that use

distinct metrics to assess a common outcome such as pain. However, these methods yield unitless measures of effect, which are not useful for estimating absolute differences.

The results of the exploratory network meta-analysis did not show a significant difference in WOMAC Pain reduction for the three comparisons of interest: NSAIDs vs. less potent opioids, NSAIDs vs. potent opioids, less potent vs. potent opioids. These results should be interpreted with caution, as there were no direct comparisons between potent opioids and either less potent opioids or NSAIDs, and there was only one indirect comparison through placebo. However, there was a trend for NSAIDs to have a larger WOMAC Pain change than opioids. This finding warrants future investigation, particularly of the consistency assumption implicit in NMA that states that direct and indirect evidence must be in agreement. This assumption could be threatened by differences in populations, treatments, and outcome ascertainment.[38] Additionally, we found that placebo effects may be greater in studies evaluating opioids. Further studies should examine the consistency of the oral placebo effects in studies of pharmacologic regimens with hypothesized differential analgesic potency.

These analyses offer important implications for research, policy, and clinical care. Studies assessing the comparative effectiveness of opioids and NSAIDs are central to clarifying the proper role of these agents for chronic OA pain. While there have been various randomized controlled trials evaluating the efficacy of analgesics, reporting has not been standardized, producing literature that is difficult to compare. Although long-term randomized controlled trials comparing effectiveness of these analgesics remain the gold standard, such studies are presently unavailable. Our results suggest that opioids provide similar levels of analgesia as NSAIDs; moreover, similar pain relief is observed for less potent and potent opioids. In addition to giving clinicians a practical way to consider the effectiveness of these analgesics, the results we present can also be used in decision analysis modeling to help policy-makers understand the role of these analgesics in the treatment of knee OA and prioritize future data acquisition.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors wish to thank Heidi Y. Yang, MS, MPH, for her assistance with statistical analyses.

Funding: This study was supported by grants from the National Institutes of Health / National Institute of Arthritis and Musculoskeletal and Skin Diseases: R01 AR064320 (Losina), K24 AR057827 (Losina), and P60 AR047782 (Katz).

Role of the Sponsors: The funding organization had no role in the design or conduct of the study, the collection, management, analysis, or interpretation of data, the preparation, review, or approval of the manuscript, or the decision to submit the manuscript for publication.

References

1. Losina E, Weinstein AM, Reichmann WM, Burbine SA, Solomon DH, Daigle ME, et al. Lifetime risk and age at diagnosis of symptomatic knee osteoarthritis in the US. Arthritis Care Res. 2013; 65:703–711.
2. Losina E, Paltiel AD, Weinstein AM, Yelin E, Hunter DJ, Chen SP, et al. Lifetime medical costs of knee osteoarthritis management in the United States: impact of extending indications for total knee arthroplasty. Arthritis Care Res. 2015; 67:203–215.
3. Jevsevar DS, Brown GA, Jones DL, Matzkin EG, Manner PA, Mooar P, et al. The American Academy of Orthopaedic Surgeons evidence-based guideline on: treatment of osteoarthritis of the knee, 2nd edition. J Bone Joint Surg Am. 2013; 95:1885–1886. [PubMed: 24288804]
4. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, et al. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. Arthritis Care Res. 2012; 64:465–474.
5. Jordan KM, Arden NK, Doherty M, Bannwarth B, Bijlsma JW, Dieppe P, et al. EULAR Recommendations 2003: an evidence based approach to the management of knee osteoarthritis: Report of a Task Force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). Ann Rheum Dis. 2003; 62:1145–1155. [PubMed: 14644851]
6. McAlindon TE, Bannuru RR, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis Cartilage. 2014; 22:363–388. [PubMed: 24462672]
7. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009; 6:e1000100. [PubMed: 19621070]
8. Wright EA, Katz JN, Abrams S, Solomon DH, Losina E. Trends in prescription of opioids from 2003–2009 in persons with knee osteoarthritis. Arthritis Care Res. 2014; 66:1489–1495.
9. Ray WA, Murray KT, Kawai V, Graham DJ, Cooper WO, Hall K, et al. Propoxyphene and the risk of out-of-hospital death. Pharmacoepidemiol Drug Saf. 2013; 22:403–412. [PubMed: 23408551]
10. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol. 1988; 15:1833–1840. [PubMed: 3068365]
11. Meissner K, Fassler M, Rucker G, Kleijnen J, Hrobjartsson A, Schneider A, et al. Differential effectiveness of placebo treatments: a systematic review of migraine prophylaxis. JAMA Intern Med. 2013; 173:1941–1951. [PubMed: 24126676]
12. Bannuru RR, McAlindon TE, Sullivan MC, Wong JB, Kent DM, Schmid CH. Effectiveness and Implications of Alternative Placebo Treatments: A Systematic Review and Network Meta-analysis of Osteoarthritis Trials. Ann Intern Med. 2015; 163:365–372. [PubMed: 26215539]
13. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials. 1996; 17:1–12. [PubMed: 8721797]
14. Bolognese JA, Schnitzer TJ, Ehrich EW. Response relationship of VAS and Likert scales in osteoarthritis efficacy measurement. Osteoarthritis Cartilage. 2003; 11:499–507. [PubMed: 12814613]
15. Ross, SM. Introduction to Probability Methods. Ninth. Orlando, FL, USA: Academic Press, Inc.; 2006.
16. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997; 315:629–634. [PubMed: 9310563]
17. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics. 1994; 50:1088–1101. [PubMed: 7786990]
18. Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics. 2000; 56:455–463. [PubMed: 10877304]
19. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002; 21:1539–1558. [PubMed: 12111919]

20. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. Stat Med. 2002; 21:589–624. [PubMed: 11836738]
21. Jones B, Roger J, Lane PW, Lawton A, Fletcher C, Cappelleri JC, et al. Statistical approaches for conducting network meta-analysis in drug development. Pharm Stat. 2011; 10:523–531. [PubMed: 22213533]
22. Mills EJ, Thorlund K, Ioannidis JP. Demystifying trial networks and network meta-analysis. BMJ. 2013; 346:f2914. [PubMed: 23674332]
23. Hochberg MC, Martel-Pelletier J, Monfort J, Moller I, Castillo JR, Arden N, et al. Combined chondroitin sulfate and glucosamine for painful knee osteoarthritis: a multicentre, randomised, double-blind, non-inferiority trial versus celecoxib. Ann Rheum Dis. 2016; 75:37–44. [PubMed: 25589511]
24. DeLemos BP, Xiang J, Benson C, Gana TJ, Pascual ML, Rosanna R, et al. Tramadol hydrochloride extended-release once-daily in the treatment of osteoarthritis of the knee and/or hip: a double-blind, randomized, dose-ranging trial. Am J Ther. 2011; 18:216–226. [PubMed: 20215961]
25. Emkey R, Rosenthal N, Wu SC, Jordan D, Kamin M. Efficacy and safety of tramadol/acetaminophen tablets (Ultracet) as add-on therapy for osteoarthritis pain in subjects receiving a COX-2 nonsteroidal antiinflammatory drug: a multicenter, randomized, double-blind, placebo-controlled trial. J Rheumatol. 2004; 31:150–156. [PubMed: 14705234]
26. Fishman RL, Kistler CJ, Ellerbusch MT, Aparicio RT, Swami SS, Shirley ME, et al. Efficacy and safety of 12 weeks of osteoarthritic pain therapy with once-daily tramadol (Tramadol Contramid OAD). J Opioid Manag. 2007; 3:273–280. [PubMed: 18181382]
27. Verkleij SP, Luijsterburg PA, Bohnen AM, Koes BW, Bierma-Zeinstra SM. NSAIDs vs acetaminophen in knee and hip osteoarthritis: a systematic review regarding heterogeneity influencing the outcomes. Osteoarthritis Cartilage. 2011; 19:921–929. [PubMed: 21619937]
28. Bjordal JM, Klovning A, Ljunggren AE, Slordal L. Short-term efficacy of pharmacotherapeutic interventions in osteoarthritic knee pain: A meta-analysis of randomised placebo-controlled trials. Eur J Pain. 2007; 11:125–138. [PubMed: 16682240]
29. Myers J, Wielage RC, Han B, Price K, Gahn J, Paget MA, et al. The efficacy of duloxetine, non-steroidal anti-inflammatory drugs, and opioids in osteoarthritis: a systematic literature review and meta-analysis. BMC Musculoskelet Disord. 2014; 15:76. [PubMed: 24618328]
30. Bannuru RR, Schmid CH, Kent DM, Vaysbrot EE, Wong JB, McAlindon TE. Comparative effectiveness of pharmacologic interventions for knee osteoarthritis: a systematic review and network meta-analysis. Ann Intern Med. 2015; 162:46–54. [PubMed: 25560713]
31. Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. Int J Technol Assess Health Care. 2012; 28:138–144. [PubMed: 22559755]
32. Bates MS, Edwards WT, Anderson KO. Ethnocultural influences on variation in chronic pain perception. Pain. 1993; 52:101–112. [PubMed: 8446431]
33. Callister LC. Cultural influences on pain perceptions and behaviors. Home Health Care Manag Pract. 2003; 15:207–211.
34. Klinge SA, Sawyer GA. Effectiveness and safety of topical versus oral nonsteroidal anti-inflammatory drugs: a comprehensive review. Phys Sportsmed. 2013; 41:64–74. [PubMed: 23703519]
35. Argoff CE. Topical analgesics in the management of acute and chronic pain. Mayo Clin Proc. 2013; 88:195–205. [PubMed: 23374622]
36. Sterne JA, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. BMJ. 2001; 323:101–105. [PubMed: 11451790]
37. Bellamy N, Campbell J, Syrotuik J. Comparative study of self-rating pain scales in osteoarthritis patients. Curr Med Res Opin. 1999; 15:113–119. [PubMed: 10494494]
38. Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. J Am Stat Assoc. 2006; 101:447–459.
39. Bingham CO 3rd, Sebba AI, Rubin BR, Ruoff GE, Kremer J, Bird S, et al. Efficacy and safety of etoricoxib 30 mg and celecoxib 200 mg in the treatment of osteoarthritis in two identically

designed, randomized, placebo-controlled, non-inferiority studies. Rheumatology. 2007; 46:496–507. [PubMed: 16936327]

40. Conaghan PG, Dickson J, Bolten W, Cevc G, Rother M. A multicentre, randomized, placebo- and active-controlled trial comparing the efficacy and safety of topical ketoprofen in Transfersome gel (IDEA-033) with ketoprofen-free vehicle (TDT 064) and oral celecoxib for knee pain associated with osteoarthritis. Rheumatology. 2013; 52:1303–1312. [PubMed: 23542612]
41. Fleischmann R, Sheldon E, Maldonado-Cocco J, Dutta D, Yu S, Sloan VS. Lumiracoxib is effective in the treatment of osteoarthritis of the knee: a prospective randomized 13-week study versus placebo and celecoxib. Clin Rheumatol. 2006; 25:42–53. [PubMed: 16132165]
42. Lehmann R, Brzosko M, Kopsa P, Nischik R, Kreisse A, Thurston H, et al. Efficacy and tolerability of lumiracoxib 100 mg once daily in knee osteoarthritis: a 13-week, randomized, double-blind study vs. placebo and celecoxib. Curr Med Res Opin. 2005; 21:517–526. [PubMed: 15899100]
43. Sheldon E, Beaulieu A, Paster Z, Dutta D, Yu S, Sloan VS. Efficacy and tolerability of lumiracoxib in the treatment of osteoarthritis of the knee: a 13-week, randomized, double-blind comparison with celecoxib and placebo. Clin Ther. 2005; 27:64–77. [PubMed: 15763607]
44. Tannenbaum H, Berenbaum F, Reginster JY, Zacher J, Robinson J, Poor G, et al. Lumiracoxib is effective in the treatment of osteoarthritis of the knee: a 13 week, randomised, double blind study versus placebo and celecoxib. Ann Rheum Dis. 2004; 63:1419–1426. [PubMed: 15020310]
45. Case JP, Baliunas AJ, Block JA. Lack of efficacy of acetaminophen in treating symptomatic knee osteoarthritis: a randomized, double-blind, placebo-controlled comparison trial with diclofenac sodium. Arch Intern Med. 2003; 163:169–178. [PubMed: 12546607]
46. Kriegel W, Korff KJ, Ehrlich JC, Lehnhardt K, Macciocchi A, Moresino C, et al. Double-blind study comparing the long-term efficacy of the COX-2 inhibitor nimesulide and naproxen in patients with osteoarthritis. Int J Clin Pract. 2001; 55:510–514. [PubMed: 11695069]
47. Raynauld JP, Martel-Pelletier J, Bias P, Laufer S, Haraoui B, Choquette D, et al. Protective effects of licofelone, a 5-lipoxygenase and cyclo-oxygenase inhibitor, versus naproxen on cartilage loss in knee osteoarthritis: a first multicentre clinical trial using quantitative MRI. Ann Rheum Dis. 2009; 68:938–947. [PubMed: 18653484]
48. Ayral X, Mackillop N, Genant HK, Kirkpatrick J, Beaulieu A, Pippingskiold P, et al. Arthroscopic evaluation of potential structure-modifying drug in osteoarthritis of the knee. A multicenter, randomized, double-blind comparison of tenidap sodium vs piroxicam. Osteoarthritis Cartilage. 2003; 11:198–207. [PubMed: 12623291]
49. Gana TJ, Pascual ML, Fleming RR, Schein JR, Janagap CC, Xiang J, et al. Extended-release tramadol in the treatment of osteoarthritis: a multicenter, randomized, double-blind, placebo-controlled clinical trial. Curr Med Res Opin. 2006; 22:1391–1401. [PubMed: 16834838]
50. Hale M, Tudor IC, Khanna S, Thipphawong J. Efficacy and tolerability of once-daily OROS hydromorphone and twice-daily extended-release oxycodone in patients with chronic, moderate to severe osteoarthritis pain: results of a 6-week, randomized, open-label, noninferiority analysis. Clin Ther. 2007; 29:874–888. [PubMed: 17697906]
51. Vojtassak J, Jacobs A, Rynn L, Waechter S, Richarz U. A phase IIIb, multicentre, randomised, parallel-group, placebo-controlled, double-blind study to investigate the efficacy and safety of OROS hydromorphone in subjects with moderate-to-severe chronic pain induced by osteoarthritis of the hip or the knee. Pain Res Treat. 2011; 2011:239501. [PubMed: 22110921]

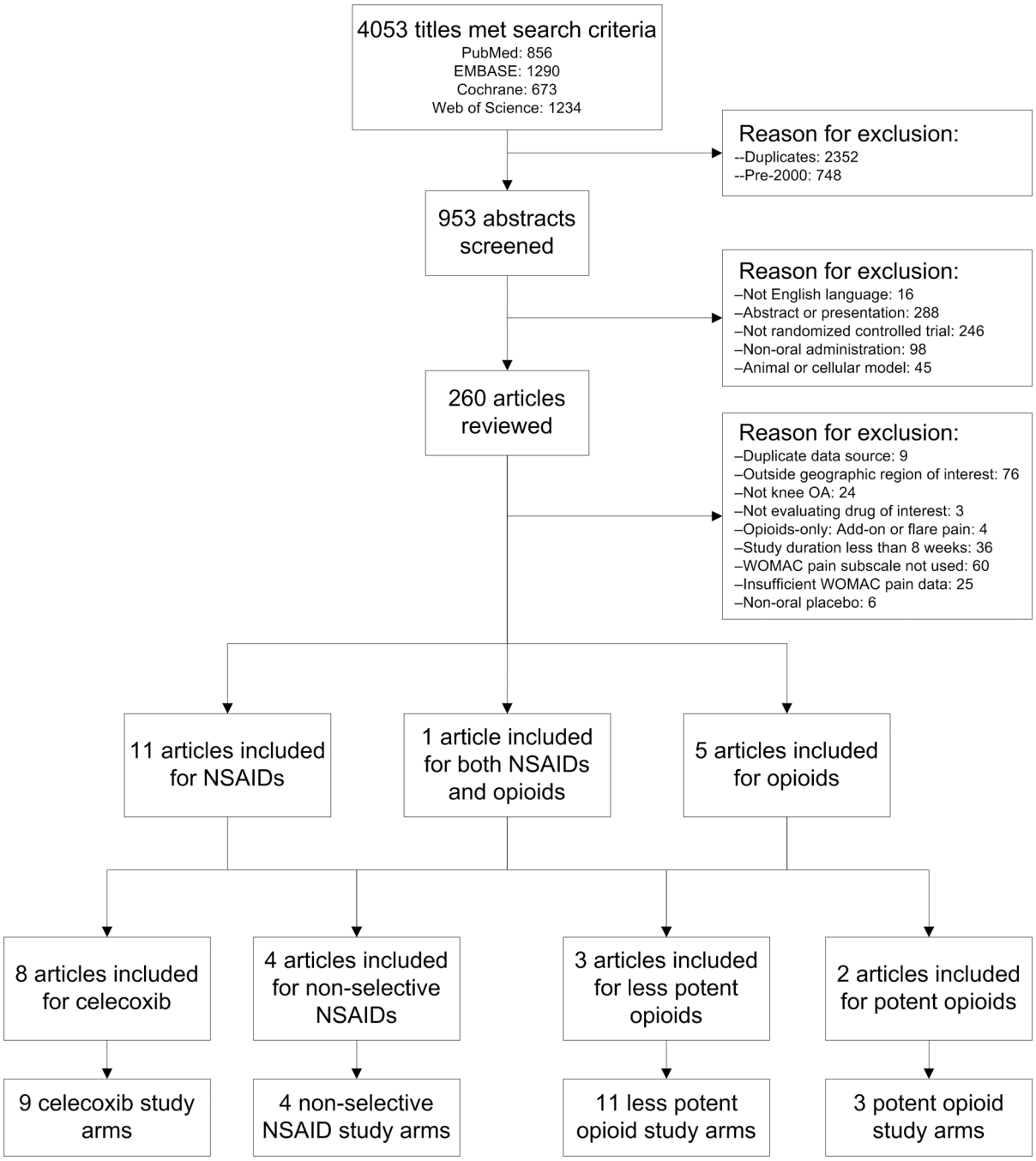
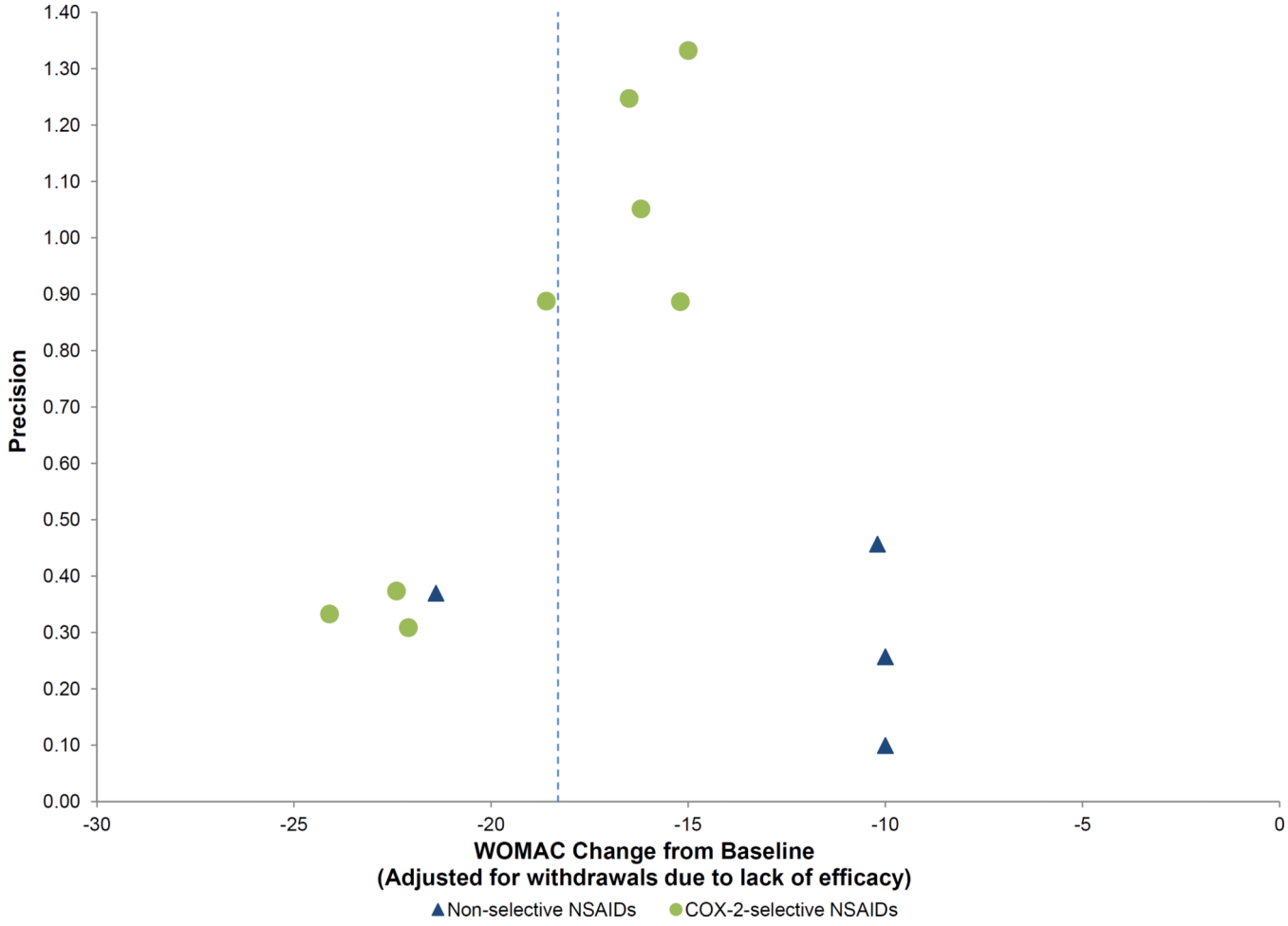


Figure 1.

One study included both NSAIDs and opioids treatment groups and is represented in both arms in this figure. Abbreviations: OA, osteoarthritis; NSAID, non-steroidal anti-inflammatory drug; WOMAC Pain, Western Ontario and McMaster Universities Osteoarthritis Index Pain Subscale



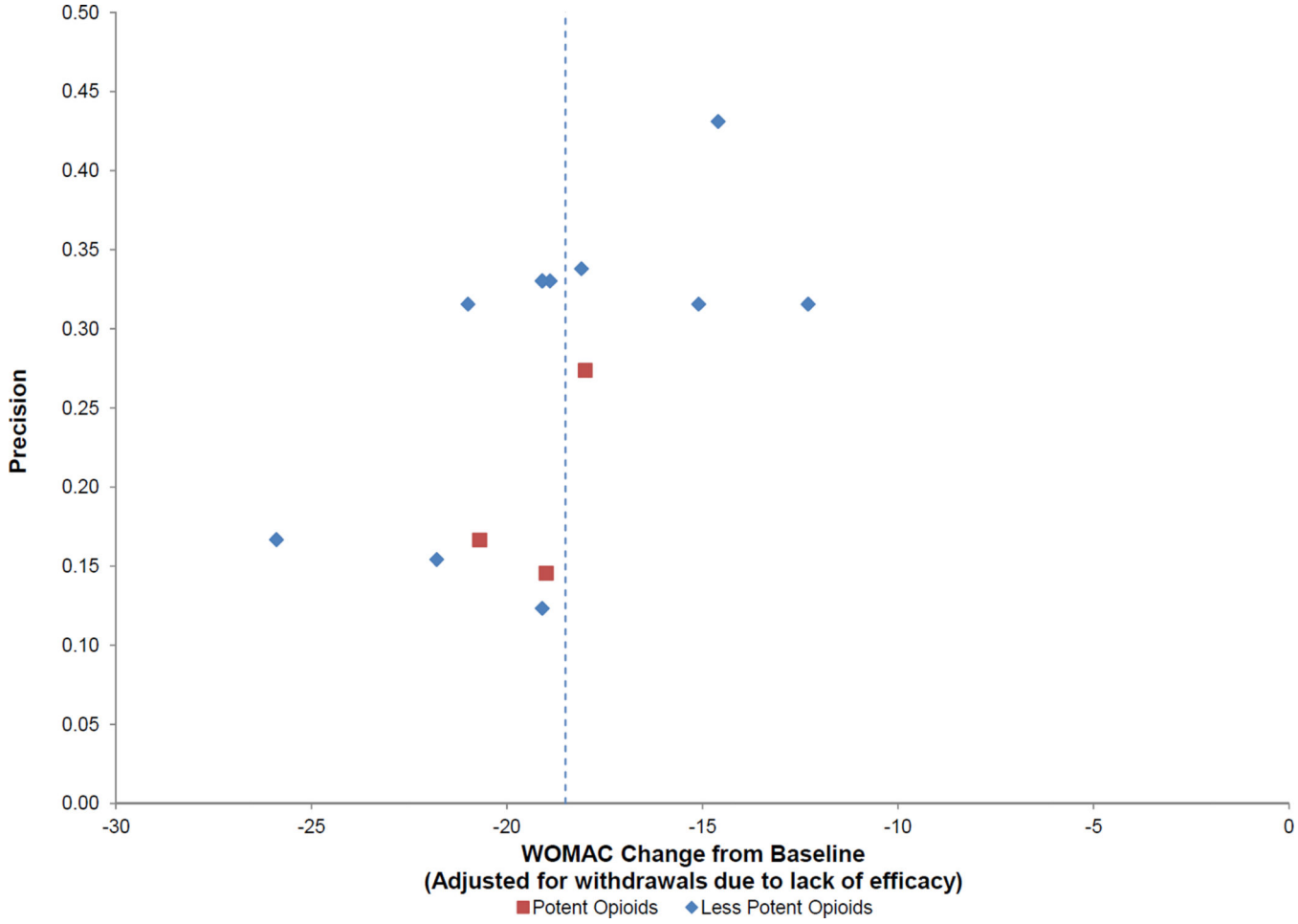


Figure 2.

This figure displays the funnel plot of precision (reciprocal of the variance) by mean change from baseline, modified for efficacy-related withdrawals, in WOMAC Pain for all included studies of (a) NSAIDs and (b) opioids. The NSAIDs funnel plot appears fairly symmetrical, and Egger's test was not statistically significant (p=0.50). The opioids funnel plot is asymmetrical, with more studies reporting more change and fewer studies with lower precision reporting less change; Egger's test was borderline statistically significant (p=0.05). The dashed lines represent combined efficacy estimates from a random effects model of the two classes of analgesics (NSAIDs: −18; opioids: −19).

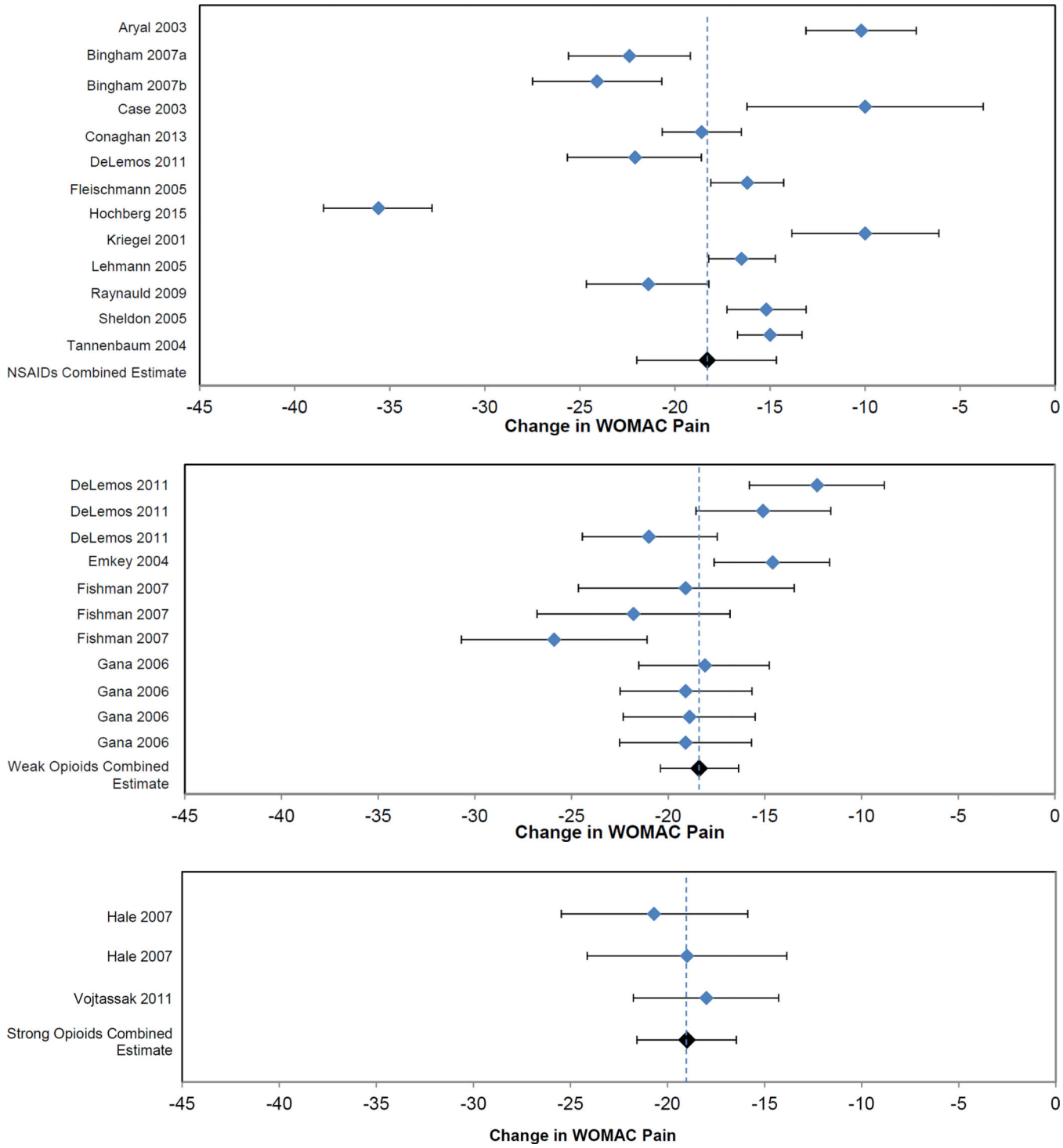


Figure 3.

This figure portrays the mean change from baseline, modified for efficacy-related withdrawals, in WOMAC Pain (with 95% confidence intervals [CIs]) for all included studies of (a) NSAIDs, (b) less potent opioids, and (c) potent opioids.

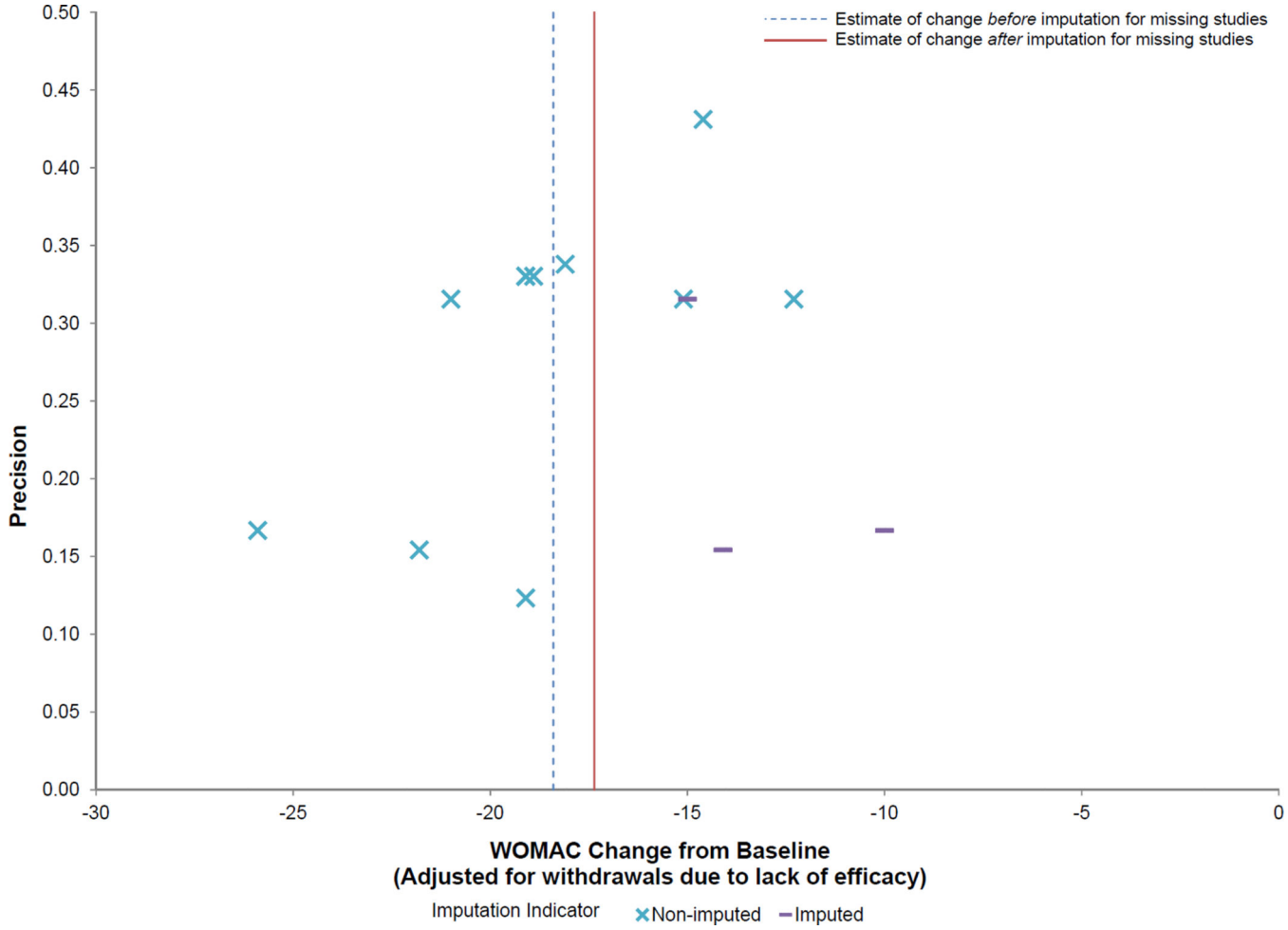


Figure 4.

The trim and fill method was utilized to impute hypothetical missing publications (indicated with a dash) for less potent opioids, as the funnel plot initially exhibited significant asymmetry. Though the peak remains uncentered, the plot is more symmetric, and Egger's test is no longer statistically significant (p=0.57).

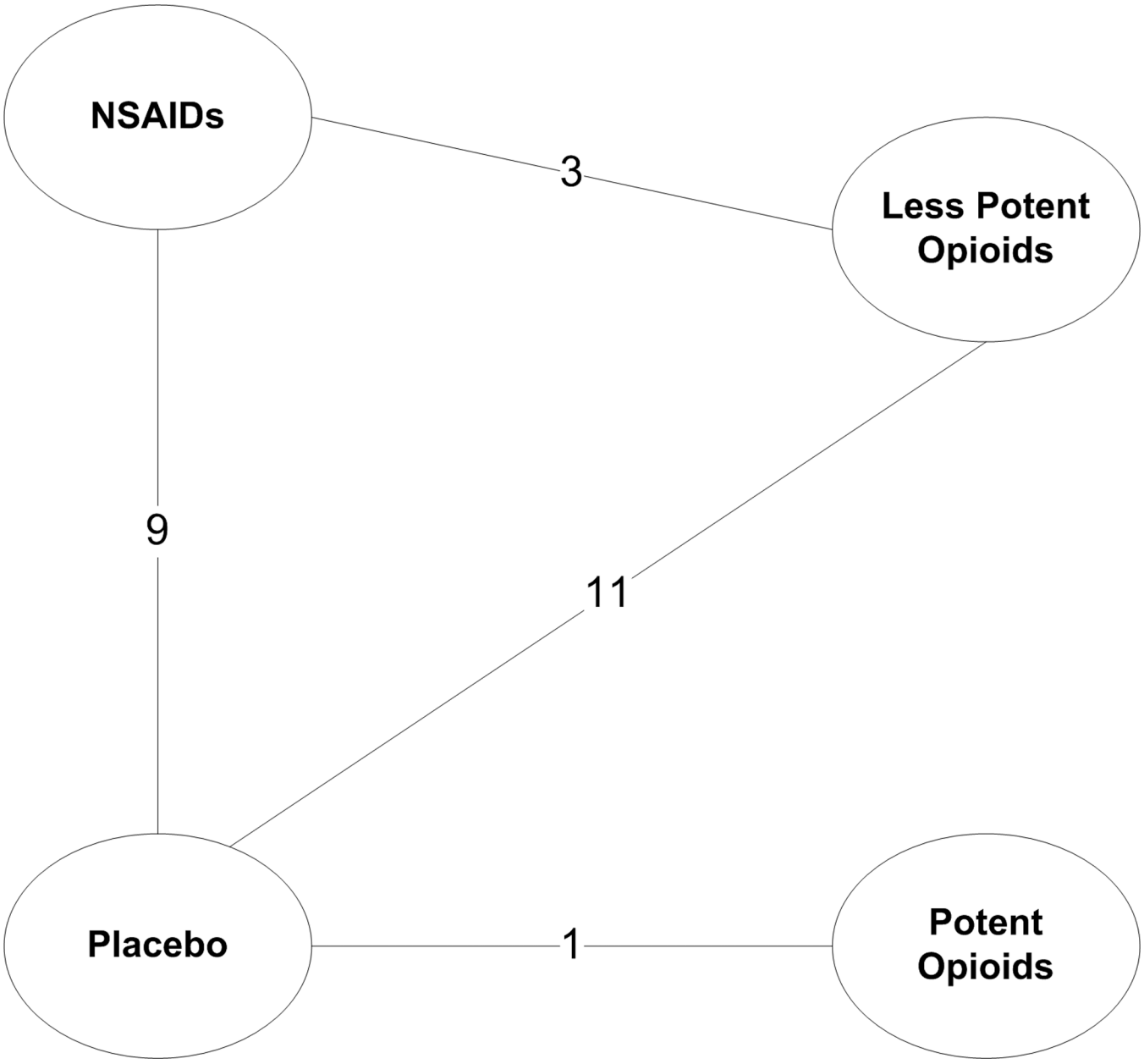


Figure 5.

This figure depicts the direct and indirect comparisons between NSAIDs, less potent opioids, potent opioids, and placebo treatment arms among included studies. Abbreviations: NSAIDs, non-steroidal anti-inflammatory drugs.

Table 1

Characteristics of included studies

Study	Analgesic evaluated (daily total dose)	n ITT	Mean age (years)	% female	Mean BMI	Mean OA duration (years)	% knee OA	Baseline WOMAC Pain (Mean, SD)	Study duration (weeks)	Jadad score
NSAIDs										
Bingham *et al.* 2007a[39]	Celecoxib (200mg)	236	63	70			80	68 (16)	12	4
Bingham *et al.* 2007b[39]	Celecoxib (200mg)	246	62	62			78	67 (19)	12	4
Conaghan *et al.* 2013[40]	Celecoxib (200mg)	233	62	67			100	47 (10)	12	3
DeLemos *et al.* 2011[24]	Celecoxib (200mg)	202	60	65		8	74	57 (19)	12	4
Fleischmann *et al* 2006[41]	Celecoxib (200mg)	444	61	67	32	7	100	52 (17)	13	4
Hochberg *et al.* 2016[23]	Celecoxib (200mg)	282[a]	63	81	31		100	74 (8)	26	5
Lehmann *et al.* 2005[42]	Celecoxib (200mg)	420	63	68	30	4	100	51 (16)	13	5
Sheldon *et al.* 2005[43]	Celecoxib (200mg)	393	60	63	33	7	100	54 (16)	13	3
Tannenbaum *et al.* 2004[44]	Celecoxib (200mg)	481	64	69	30	5	100	51 (17)	13	4
Case *et al.* 2003[45]	Diclofenac (150mg)	25[a]	63	60	27		100	40 (20)	12	4
Kriegel *et al.* 2001[46]	Naproxen (750mg)	187	65	75		4	71	48 (19)	26	4
Raynauld *et al.* 2009[47]	Naproxen (1000mg)	154	60	65	31		100	58 (12)	26	4
Aryal *et al.* 2003[48]	Piroxicam (20mg)	221	61	61	31	3	100	40 (19)	52	3
Less Potent Opioids										
DeLemos *et al.* 2011[24]	Tramadol (100mg)	201	60	58		8	74	60 (20)	12	4
DeLemos *et al.* 2011[24]	Tramadol (200mg)	199	62	62		9	73	61 (19)	12	4
DeLemos *et al.* 2011[24]	Tramadol (300mg)	199	60	62		8	75	61 (21)	12	4

Study	Analgesic evaluated (daily total dose)	n ITT	Mean age (years)	% female	Mean BMI	Mean OA duration (years)	% knee OA	Baseline WOMAC Pain (Mean, SD)	Study duration (weeks)	Jadad score
Fishman *et al.* 2007[26]	Tramadol (100mg)	103	63	60	31		100	58 (16)	12	5
Fishman *et al.* 2007[26]	Tramadol (200mg)	107	61	60	30		100	57 (16)	12	5
Fishman *et al.* 2007[26]	Tramadol (300mg)	105	60	66	31		100	63 (19)	12	5
Gana *et al.* 2006[49]	Tramadol (100mg)	202	58	62	34	8	75	62 (20)	12	5
Gana *et al.* 2006[49]	Tramadol (200mg)	201	59	64	34	8	74	63 (18)	12	5
Gana *et al.* 2006[49]	Tramadol (300mg)	201	59	59	33	8	74	59 (19)	12	5
Gana *et al.* 2006[49]	Tramadol (400mg)	202	58	58	34	8	74	60 (19)	12	5
Emkey *et al.* 2004[25]	Tramadol[b] (150–300mg)	153	60	65			82	54 (14)	13	3
Potent Opioids										
Hale *et al.* 2007[50]	Hydromorphone (8–64mg)	64	63	77	34		84	63 (16)	8	3
Vojtassak *et al.* 2011[51]	Hydromorphone (4mg)	138	65	77	32		76	59 (13)	16	5
Hale *et al.* 2007[50]	Oxycodone (20–160mg)	60	64	62	31		75	61 (15)	8	3

[a] Data for intention-to-treat population not available; total randomized population evaluated instead

[b] Tramadol in conjunction with acetaminophen (1300–2600mg)

Abbreviations: ITT, modified intention-to-treat population; BMI, body mass index; OA, osteoarthritis; OA duration, years since diagnosed with osteoarthritis; WOMAC Pain, Western Ontario and McMaster Universities Osteoarthritis Index pain subscale; SD, standard deviation

Table 2

Mean change and adjusted mean change in WOMAC Pain for included studies

Study	Analgesic evaluated	Mean Change (SD)	Adjusted Mean Change[a]
NSAIDs			
Bingham *et al.* 2007a[39]	Celecoxib (200mg)	−25 (25)	−22
Bingham *et al.* 2007b[39]	Celecoxib (200mg)	−27 (27)	−24
Conaghan *et al.* 2013[40]	Celecoxib (200mg)	−19 (16)	−19
DeLemos *et al.* 2011[24]	Celecoxib (200mg)	−26 (26)	−22
Fleischmann *et al.* 2006[41]	Celecoxib (200mg)	−18 (21)	−16
Hochberg *et al.* 2016[23]	Celecoxib (200mg)	−37 (24)	−36
Lehmann *et al.* 2005[42]	Celecoxib (200mg)	−17 (18)	−17
Sheldon *et al.* 2005	Celecoxib (200mg)	−17 (21)	−15
Tannenbaum *et al.* 2004[44]	Celecoxib (200mg)	−16 (19)	−15
Case *et al.* 2003[45]	Diclofenac (150mg)	−11 (16)	−10
Kriegel *et al.* 2001[46]	Naproxen (750mg)	−11 (27)	−10
Raynauld *et al.* 2009[47]	Naproxen (1000mg)	−25 (20)	−21
Aryal *et al.* 2003[48]	Piroxicam (20mg)	−11 (22)	−10
Less Potent Opioids			
DeLemos *et al.* 2011[24]	Tramadol (100mg)	−17 (25)	−12
DeLemos *et al.* 2011[24]	Tramadol (200mg)	−18 (25)	−15
DeLemos *et al.* 2011[24]	Tramadol (300mg)	−24 (25)	−21
Fishman *et al.* 2007[26]	Tramadol (100mg)	−24 (29)	−19
Fishman *et al.* 2007[26]	Tramadol (200mg)	−24 (26)	−22
Fishman *et al.* 2007[26]	Tramadol (300mg)	−29 (25)	−26
Gana *et al.* 2006[49]	Tramadol (100mg)	−21 (24)	−18
Gana *et al.* 2006[49]	Tramadol (200mg)	−22 (25)	−19
Gana *et al.* 2006[49]	Tramadol (300mg)	−21 (25)	−19
Gana *et al.* 2006[49]	Tramadol (400mg)	−22 (25)	−19
Emkey *et al.* 2004[25]	Tramadol (150–300mg)[b]	−16 (19)	−15
Potent Opioids			
Hale *et al.* 2007[50]	Hydromorphone (8–64mg)	−21 (20)	−21
Vojtassak *et al.* 2011[51]	Hydromorphone (4mg)	−19 (22)	−18
Hale *et al.* 2007[50]	Oxycodone (20–160mg)	−20 (20)	−19

[a] Modified for withdrawals due to insufficient efficacy

[b] Tramadol in conjunction with acetaminophen (1300–2600mg)

Abbreviations: WOMAC Pain, Western Ontario and McMaster Universities Osteoarthritis Index pain subscale; SD, standard deviation