Published in final edited form as:

Qual Life Res. 2017 August ; 26(8): 1955–1967. doi:10.1007/s11136-017-1538-0.

Health-related quality of life in patients receiving long-term opioid therapy: a systematic review with meta-analysis

J. Douglas Thornton¹, Rashmi Goyat¹, Nilanjana Dwibedi¹, and George A. Kelley²

¹Department of Pharmaceutical Systems and Policy, West Virginia University School of Pharmacy, Robert C. Byrd Health Sciences Center (North), P.O. Box 9510, Morgantown, WV 26506-9510, USA

²Department of Biostatistics, School of Public Health, West Virginia University, Morgantown, WV, USA

Abstract

Purpose—Over 25 million Americans reported having daily pain and between 5 and 8 million Americans used opioids to treat chronic pain in 2012. This is the first systematic review with meta-analysis to determine the effects of long-term opioid use on the Physical Component Summary (PCS) score and Mental Component Summary (MCS) scores of a Health-Related Quality of Life instrument in adults without opioid use disorder.

Methods—The a priori eligibility criteria for the PubMed (MEDLINE), Scopus, and PsycINFO searches were (1) randomized controlled trial, (2) at least one opioid intervention group, (3) minimum of 4-week duration of opioid use, (4) comparative control group, and (5) adults 18 years that do not have dominant disease. The unit of analysis was the standardized mean difference effect size (Hedges's g). All results were pooled using random-effects models.

Results—Of the 340 non-duplicate citations screened, 19 articles comprising 26 treatment comparisons and 6168 individuals (treatment $n = 3160$; comparators $n = 3008$ with duplicates removed) met the inclusion criteria for the systematic review. Thirteen treatment comparisons were available for the meta-analysis. Across all PCS analyses, small, statistically significant improvements were observed (opioid versus opioid only: $g = 0.27$, 95% CI 0.05–0.50, opioid versus placebo only: $g = 0.18$, 95% CI 0.08–0.28, and all studies combined: $g = 0.22$, 95% CI 0.11–0.32). There were small but not statistically significant changes on the MCS scores. Overall, high heterogeneity was present.

Conclusions—PCS scores improve with no change in MCS scores. However, long-term opioid trials are rare and only two trials included lasted longer than 1 year.

Correspondence to: Nilanjana Dwibedi.

Electronic supplementary material The online version of this article (doi:10.1007/s11136-017-1538-0) contains supplementary material, which is available to authorized users.

Compliance with ethical standards

Conflict of interest JDT, RG, ND, and GAK declare that they have no conflicts of interests.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Keywords

Health-related quality of life; Opioid; Systematic review; Meta-analysis; Randomized controlled trials

Introduction

Chronic pain has been defined as pain beyond the time of normal tissue healing, or 3 months, and is a significant cause of disability and reduced quality of life [1, 2]. The Institute of Medicine estimated in 2011 that over 100 million people in the United States experienced chronic pain [3], with costs estimated up to \$630 billion annually [4]. A more recent evaluation of the National Health Interview Study reported that 11.2% of adults in 2012 had daily pain [5]. Chronic pain as a public health problem has wide reaching consequences for families as well as economic burden relating to lost productivity [6]. An effective treatment is needed to alleviate the impact of chronic pain. Opioids are very effective in treating moderate to severe pain in the short term by reducing the pain signals reaching the brain [7] and are a generally accepted treatment for cancer-related pain [6, 8]. However, there is a lack of evidence supporting the effectiveness of opioids in treating non-cancer chronic pain [8, 9]. Despite this lack of support, in 2012, there were 259 million prescriptions written for opioid analgesics in the United States [10], and it was estimated that between 5 and 8 million Americans used opioids to treat chronic pain. Total prescribing increased by nearly 50% over the decade, and as the number of prescriptions for opioids has skyrocketed, prescription opioid misuse and abuse have increased in parallel [11].

Treating chronic pain from a prescriber's perspective can be a complicated endeavor. Some physicians have reported erring on the side of over-treating with opioid therapy, at least in the short term, because pain can impede therapeutic success in other areas [12]. Primary care providers have been issued special guidelines to emphasize appropriate opioid prescribing for patients with chronic pain [8]. However, the trends of over-prescribing should not mask the benefits that opioids can have for patients with chronic pain [13]. A nuanced approach is needed to balance effective pain management and reduction in harms from long-term opioid therapy.

One way to potentially curb inappropriate use of opioids is improved monitoring practices in patients receiving long-term opioid therapy. An appropriate monitoring strategy will regularly assess the patients' Health-related quality of life (HRQoL). Health-related quality of life focuses on aspects of life that can affect either physical or mental health [14]. It is important to monitor HRQoL so that patients can be compared across disease states, disciplines, and medical services [14]. The Medical Outcome Study Short Form-36 (SF-36), shorter versions including the SF-12 and SF-8, and disease-specific instruments have been used to measure HRQoL in patients using opioids. The SF-36 is a commonly used instrument that has been highly rated in terms of its conceptual and measurement model, reliability, validity, and administrative burden [15].

Healthcare providers use a wide range of information of varying methodological quality to make clinical decisions [16]. A systematic review uses a structured process to compile and

appraise research on a particular area or disease so that results can be compared across studies [17]. By performing a meta-analysis in addition to the systematic review, one pooled effect size can be estimated by combining and weighting all included studies' effect sizes [17]. Pooled results from high-quality meta-analyses can provide information that is more easily interpreted by clinicians, patients, researchers, and policy makers [18].

To the best of the authors' knowledge, there have been three previous systematic reviews with and without meta-analysis on HRQoL in opioid users [19–21]. The first was a systematic review limited to a specific group of opioid users, namely, patients with opioid dependence, a unique subgroup of all opioid users that needs to be evaluated separately [19]. A second systematic review that included a meta-analysis was limited to the effects of opioids on breathlessness in patients with chronic-obstructive pulmonary disease (COPD) [20]. While it proposed to assess HRQoL, no outcomes derived from commonly used instruments for the assessment of HRQoL were included [20]. The third systematic review, limited to comparing a desired intervention (tapentadol) to one comparator (oxycodone/naloxone), did not assess HRQoL [21]. Given the lack of consolidated evidence regarding the effects of long-term opioid use on HRQoL, the objective of this study was to conduct a systematic review with meta-analysis to determine the effects of long-term opioid use on HRQoL based on Physical Component Summary (PCS) scores and Mental Component Summary (MCS) scores in adults with chronic pain and no opioid use disorder.

Methods

Overview

This study followed the general guidelines of the Preferred reporting items for systematic reviews and meta-analysis (PRISMA) of health care interventions [22].

Study eligibility

The a priori inclusion criteria for this systematic review with meta-analysis was as follows: (1) randomized controlled trial with the unit of analysis being the treatment arm, (2) at least one opioid intervention group, (3) minimum of 4-week duration of opioid use, (4) comparative control group, (5) adults 18 years of age and older that do not have a dominant disease state (chronic kidney disease, HIV/AIDs, or cancer), and (6) studies indexed in any language before April 1, 2016. The decision was made to include randomized controlled trials only because HRQoL can be very disease dependent. Thus, randomization is probably the only way to effectively account for this variation. Chronic pain was defined as beyond the time of normal tissue healing or 3 months, a time during which opioid therapy is recommended. Long-term opioid therapy was classified as at least 4 weeks of use for study inclusion. This relatively short time frame was chosen in order to include more studies since so few RCTs lasted longer than 3 months (see Table 1). Opioid use included opioid agonists as well as buprenorphine (an opioid partial agonist) when they were being used to treat chronic pain in patients that did not have opioid use disorder or similar diagnosis. While studies were not limited to full articles, there had to be enough information available to assess treatment, comparator, length of follow-up, and HRQoL assessment outcome.

Data sources

Studies were retrieved using the following three electronic databases: (1) PubMed (MEDLINE), (2) Scopus, and (3) PsycINFO. All electronic searches were performed by the first and second authors (DT and RG). While the specific search strategies varied by database slightly, the following is an example that was used for PubMed (MEDLINE):

("health"[MeSH Terms] OR "health"[All Fields]) AND related[All Fields] AND quality[All Fields] AND ("life"[MeSH Terms] OR "life"[All Fields] OR "SF*" OR "EQ-5D") AND ("analgesics, opioid"[Pharmacological Action] OR "analgesics, opioid"[MeSH Terms] OR ("analgesics"[All Fields] AND "opioid"[All Fields]) OR "opioid analgesics"[All Fields] OR "opioid"[All Fields]) AND Random*.

The Number Needed to Read (NNR) was calculated as the inverse of the included studies divided by the total number of non-duplicate citations. All studies were stored in EndNote X7 (© Thomson Reuters) and then exported to Microsoft Excel (2013) for exclusion criteria to be applied.

Study selection

Studies were selected by the first and second authors independently (DT and RG). Multiple publication bias was handled by excluding additional studies that analyzed the same patient pool. If two studies compared the same patients, the manuscript that had the most complete results on the primary outcome (HRQoL measures) was used. If an individual study included three or more groups (at least two of which used opioid medications) the control group was used as the comparator for the treatment arms. All included studies, excluded studies, and reasons for exclusion were stored in Microsoft Excel (2013).

Data abstraction

Prior to abstraction, a codebook was created that could contain 49 variables per study. The codebook was piloted and refined after discussions with the fourth author (GK). The main data categories included the following: (1) study information, (2) participant information, (3) treatment and comparator information, (4) primary outcomes (HRQoL measures), and (5) funding information. All studies were coded independently by the first and second authors (DT and RG). The third author (ND) arbitrated all disagreements. Health-related quality of life measures were included if there was enough information available to calculate an effect size and 95% confidence interval. The effect size used for this meta-analysis was the standardized mean difference (Hedge's *g*) [43].

Risk of bias assessment

The Cochrane Risk of Bias instrument was used to assess bias in six domains: (1) random sequence generation, (2) allocation concealment, (3) blinding of participants and personnel, (4) blinding of outcome assessors, (5) incomplete outcome data, and (6) selective reporting. Each study was classified as having a high, low, or unclear risk of bias. Assessment for blinding of outcome assessors was for the HRQoL assessment only since it was the primary outcome for this meta-analysis. Based on previous research, no study was excluded because

it was classified as having a “high” risk of bias [44]. All assessments were performed by the first and second authors (DT and RG).

Statistical analysis

Effect sizes were calculated for primary outcomes (HRQoL) for each study where information was available. The effect size used was the Hedge’s g standardized mean difference [43], and these were recorded in the code book. If a mean difference was not available, the Campbell Collaboration’s Practical Meta-Analysis Effects Size Calculator was used to convert available statistics to g [45]. An aggregate level meta-analysis was performed using MetaXL [46]. Pooled estimates for outcomes were calculated using random-effects modeling [47] and heterogeneity as well as inconsistency of results between studies was estimated using the Q statistic and I^2 . Inconsistency for I^2 was categorized as low (25%), moderate (50%), and high (75%) [48], while the magnitude of effect sizes were interpreted as very small (0.01), small (0.2), medium (0.5), large (0.8), very large (1.2), and huge (2.0) [49]. Small study effects were examined qualitatively using a funnel plot. Sensitivity analysis was performed by iteratively excluding one study at a time to see how the overall effect size changed. To aid practical application, the number-needed-to treat (NNT) [50] and Cohen’s U_3 index for percentile improvement [51] were calculated for those overall findings reported as statistically significant. For NNT, the method of Kraemer and Kupfer [50] was used versus a method based on control group risk given the lack of consensus regarding an appropriate control group risk for opioids and HRQoL. For Cohen’s U_3 index [51], a SMD of 0.30, for example, suggests that the opioid group would be at approximately the 62nd percentile with respect to improving their HRQoL. This would equate to being approximately 12 percentiles higher than control group participants.

Results

Study characteristics

A description of the study characteristics is shown in Table 1. There were 340 non-duplicate citations screened after 145 duplicates were removed. A total of 321 citations were removed after applying the inclusion and exclusion criteria. There were 305 removed after screening the title and abstract only, and then an additional 16 removed after reviewing the full text. A final list of 19 studies were included in the systematic review [23–29, 31–42]. One study identified using the search strategy [29] referenced a separate study [30] which was used for partial data abstraction despite not being identified via the search strategy. The NNR was 18. Eight citations contributed 13 sets of effect sizes for the meta-analysis [23, 24, 27, 29, 34, 36, 37, 39]. A flow chart showing each step of this process is shown in Fig. 1. The broad reasons for exclusion were as follows: (1) wrong population ($n = 129$), (opioid dependent/opioid use disorder, etc.), (2) inappropriate intervention ($n = 135$), [no group on opioids ($n = 110$), acute use ($n = 17$), etc.], (3) incorrect study design ($n = 50$), (non-randomized controlled trial, etc.), or (4) other ($n = 7$), (same participants as another citation, etc.). The full list of excluded studies is shown in Supplementary file (Excel sheet 2). The number of opioid treatment groups analyzed in the meta-analysis ($n = 13$) was greater than the number of studies included (8) because one study included more than one opioid treatment group. The number of unique participants varied widely within the treatment [$n = 3160$, mean (SD)

= 126.4 (83.1), median = 130, range 28–338] and comparator groups [$n = 3008$, mean (SD) = 130.8 (89.5), median = 131, range 28–342]. While all included studies were published in English, 13 of 19 (68.4%) were conducted outside the United States (China [33], Canada [28, 41], Sweden [31], Norway [36], and multisite in Europe [24, 26, 27, 29, 40, 42], North America [37], or Oceania [23]). The types of control groups varied between active controls with an opioid [24, 26, 27, 29, 31, 34–36, 38, 40, 52], active placebos with benztropine [32, 41], and inactive placebos [23, 25, 28, 33, 37, 39]. Of the 13 treatment groups included for the meta-analysis, 13 had PCS scores reported [23, 24, 27, 29, 34, 36, 37, 39] and 11 had MCS scores reported [23, 24, 29, 36, 37, 39].

Risk of bias assessment

All 19 studies included in the systematic review were evaluated for risk of bias using the Cochrane Risk of Bias assessment tool. 7 of the 19 studies were classified as having low risk of bias [25, 28, 32, 34, 36, 37, 39], and 12 were classified as having high risk of bias overall [23, 24, 26, 27, 29, 31, 33, 35, 38, 40–42]. The most common reason for a study to be classified as having a high risk of bias was not blinding participants and or personnel (open-label study). The sensitivity analysis where each study was iteratively left of the model did not affect the significance (or lack thereof) for any of the four models. The study with the biggest outlier effect size for MCS [29, 41], had high risk of bias, included the only treatment arm to have transdermal fentanyl, and was the longest duration of all included studies at 56 weeks. Even if this study was removed, the overall effect significance did not change. Risk of bias assessment results are summarized in Table 2 with notes and funding sources found in the Supplementary Files [Table 1].

Quantitative synthesis (meta-analysis)

A total of six random-effects models were used to analyze both PCS and MCS scores for (1) opioid treatment group versus opioid control group, (2) opioid treatment group versus placebo control group, and (3) the combination of all treatment groups versus control groups for all eligible studies. Across all three PCS analyses, small [49] and statistically significant improvements were observed (opioid versus opioid studies only: $g = 0.27$, 95% CI 0.05–0.50, opioid versus placebo studies only: $g = 0.18$, 95% CI 0.08–0.28, and all studies combined: $g = 0.22$, 95% CI 0.11–0.32). The NNT for each of these effect sizes were 7, 10, and 9, respectively [50]. Using Cohen's U_3 index, the percentile improvement was 10.6, 7.1, and 8.7, respectively [51]. There were small [49] and not statistically significant changes in the MCS score for the three analyses. The other results for the random-effects models (standardized mean difference effect size, pooled effect size, 95% confidence intervals, weights, I^2 , Cochran's Q , and Tau-squared) are reported in Table 3.

Forest plots and funnel plots for the two models that included all eligible studies are shown in Figs. 2, 3, 4, 5. Three studies which compared opioids versus placebo (eight treatment comparisons) had both PCS and MCS reported [23, 37, 39]. The PCS and MCS in studies that compared opioid versus placebo were assessed using separate random-effects models, and the results are also shown in Table 3. The forest plots for these two models are shown in Figs. 6 and 7. As can be seen, the random-effects models for PCS scores, comparing treatment versus control groups in all eligible studies and opioid versus placebo studies only,

were statistically significant (95% confidence interval did not cross zero), while the two for MCS scores were not. Overall, there was moderate to high inconsistency [48] as well as high heterogeneity throughout the included studies.

Discussion

The current systematic review with meta-analysis found that only the PCS findings were consistent with the current literature [8, 9]. Based on our findings, treating pain with opioids of any kind should improve HRQoL with respect to pain and physical functioning. When all ten eligible studies that measured PCS were combined, there was a significant difference between treatment and comparator groups. There were eight studies which assessed the MCS component of the SF-36. When combined, there was no statistically significant difference between groups. One possible reason for this may be related to the diversity of treatments. The PCS and MCS were both assessed in three studies (eight treatment comparisons) which compared an opioid with placebo. PCS was significantly improved for treatment groups compared to placebo.

There are implications for research to be drawn from this study. This systematic review and meta-analysis highlights the brevity of randomized controlled trials for opioids designed and marketed to be used long term as well as the limited outcome reporting when it comes to HRQoL measures. The mean duration of trials included in the systematic review was only 15.3 weeks. Also, HRQoL is deemed to be an important component of patient care, so the way it is measured and reported should be improved in future randomized clinical trials. Along those lines, the standardization of reporting for HRQoL would greatly aid comparison across studies. This standardization would mean deciding on one generic or one disease-specific instrument to measure HRQoL and also how the results are reported. Most of the eligible studies which could not be included in all PCS or MCS models were excluded because they did not report specifics about the HRQoL measures. Two studies [28, 34] reported only that “no significant differences” existed. Both were funded by Purdue Pharmaceuticals. Very few studies compared opioids and/or reported any component of HRQoL completely. This is troubling since it is difficult to interpret one component score of a generic HRQoL tool in a vacuum. Regardless, there is a tendency for authors to submit, and editors to publish, studies that yield statistically significant results.

In addition to implications for research, there were also implications for practice which could be elucidated from this study. For example, comparing opioids head-to-head is difficult. This was exacerbated by the reporting of HRQoL measures being worse in opioid versus opioid trials than in the opioid versus placebo trials. The latter offer minimal clinical benefit. Health-related quality of life measures can be used in the context of an individual patient’s care and to direct therapy changes. For future research, changes in individuals’ HRQoL measures over time may be more useful in clinical decision making.

There were both strengths and limitations in this study. First, to the best of the author’s knowledge, this is the first systematic review with meta-analysis to determine the effects of long-term opioid use on the PCS score and MCS scores of a HRQoL instrument in adults with long-term opioid therapy and no opioid use disorder. This is important given the need

to examine HRQoL in patients using opioids for chronic pain. Second, and as previously mentioned, this study provides important implications for both practice and research. Third, the pooling of findings from different studies increased statistical power for the investigative team's primary endpoint, HRQoL.

In addition to strengths, there were also potential limitations. First, there appears to be a paucity of long-term opioid trials as judged by the fact that only two included trials lasted longer than one year. As a result, long-term opioid therapy could not be assessed to any great degree. We limited our studies to randomized controlled trials because they are the only way to control for unidentified confounders as well as the fact that non-randomized controlled trials tend to overestimate effects in healthcare interventions [53, 54]. However, this most likely resulted in a shorter followup time, a potential limitation of the current study. Given the former, future meta-analytic research on this topic may want to focus on observational studies. Second, HRQoL can be affected by the underlying disease state and the pain experienced by the patient. Third, while it was more feasible to analyze PCS and MCS in opioid versus placebo trials, these are of limited clinical value. Finally, inherent in any aggregate data meta-analysis is the potential for ecological fallacy.

Conclusion

PCS scores improved with no change in MCS score. However, a need exists for long-term trials that examine changes in HRQoL among long-term opioid users since the use of these medications can impact other aspects of a patient's life and their comorbid conditions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding GAK was supported by the National Institutes of Health, National Institute of General Medical Sciences (Grant Number U54GM104942). JDT was supported by the National Institutes of Health, National Institute of General Medical Sciences (Grant Number 5T32GM081741-08). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Ballantyne JC, Shin NS. Efficacy of opioids for chronic pain: A review of the evidence. *The Clinical Journal of Pain*. 2008; 24(6):469–478. [PubMed: 18574357]
2. Eriksen J, Sjogren P, Bruera E, Ekholm O, Rasmussen NK. Critical issues on opioids in chronic non-cancer pain: An epidemiological study. *Pain*. 2006; 125(1–2):172–179. [PubMed: 16842922]
3. Institute of Medicine. *Relieving pain in America: A blue-print for transforming prevention, care, education, and research*. Washington, DC: National Academies Press; 2011.
4. Gaskin DJ, Richard P. The economic costs of pain in the United States. *The Journal of Pain: Official Journal of the American Pain Society*. 2012; 13(8):715–724. [PubMed: 22607834]
5. Nahin RL. Estimates of pain prevalence and severity in adults: United States, 2012. *The Journal of Pain: Official Journal of the American Pain Society*. 2015; 16(8):769–780. [PubMed: 26028573]
6. Rosenblum A, Marsch LA, Joseph H, Portenoy RK. Opioids and the treatment of chronic pain: Controversies, current status, and future directions. *Experimental and Clinical Psychopharmacology*. 2008; 16(5):405–416. [PubMed: 18837637]

7. National Institute of Drug Abuse. What are Opioids?. Prescription Drugs of Abuse 2014. 2016. <https://www.drugabuse.gov/publications/research-reports/prescription-drugs/opioids/what-are-opioids>
8. Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain—United States. JAMA. 2016; 315:1624–1645. [PubMed: 26977696]
9. Agency for Healthcare Research and Quality. The effectiveness and risks of long-term opioid treatment of chronic pain 2014. 2016. <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1988&pageaction=displayproduct>
10. Centers for Disease Control and Prevention. Opioid painkiller prescribing 2014. CDC Vital Signs. 2016. <http://www.cdc.gov/vitalsigns/opioid-prescribing>
11. Centers for Disease Control and Prevention. Vital signs: Overdoses of prescription opioid pain relievers—United States, 1999–2008. MMWR: Morbidity and Mortality Weekly Report. 2011; 60(43):1487–1492. [PubMed: 22048730]
12. Zgierska A, Miller M, Rabago D. Patient satisfaction, prescription drug abuse, and potential unintended consequences. JAMA: The Journal of the American Medical Association. 2012; 307(13):1377–1378. [PubMed: 22474199]
13. Ballantyne JC, Mao J. Opioid therapy for chronic pain. The New England Journal of Medicine. 2003; 349(20):1943–1953. [PubMed: 14614170]
14. Centers for Disease Control and Prevention. HRQOL Concepts 2011. 2016. <http://www.cdc.gov/hrqol/concept.htm>
15. Coons SJ, Rao S, Keininger DL, Hays RD. A comparative review of generic quality-of-life instruments. Pharmacoeconomics. 2000; 17(1):13–35. [PubMed: 10747763]
16. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: Synthesis of best evidence for clinical decisions. Annals of Internal Medicine. 1997; 126(5):376–380. [PubMed: 9054282]
17. Garg AX, Hackman D, Tonelli M. Systematic review and meta-analysis: When one study is just not enough. Clinical Journal of the American Society of Nephrology. 2008; 3(1):253–260. [PubMed: 18178786]
18. Skelly AC. Credibility matters: Mind the gap. Evid Based Spine Care Journal. 2014; 5(1):2–5.
19. De Maeyer J, Vanderplasschen W, Broekaert E. Quality of life among opiate-dependent individuals: A review of the literature. International Journal on Drug Policy. 2010; 21(5):364–380. [PubMed: 20172706]
20. Ekstrom M, Nilsson F, Abernethy AA, Currow DC. Effects of opioids on breathlessness and exercise capacity in chronic obstructive pulmonary disease. A systematic review. Annals of the American Thoracic Society. 2015; 12(7):1079–1092. [PubMed: 25803110]
21. Thakur D, Dickerson S, Kumar Bhutani M, Junor R. Impact of prolonged-release oxycodone/naloxone on outcomes affecting patients' daily functioning in comparison with extended-release tapentadol: A systematic review. Clinical Therapeutics. 2015; 37(1):212–224. [PubMed: 25592091]
22. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. BMJ. 2009; 339:b2700. [PubMed: 19622552]
23. Afilalo M, Morlion B. Efficacy of tapentadol ER for managing moderate to severe chronic pain. Pain Physician. 2013; 16(1):27–40. [PubMed: 23340531]
24. Baron R, Jansen JP, Binder A, et al. Tolerability, safety, and quality of life with tapentadol prolonged release (pr) compared with oxycodone/naloxone PR in patients with severe chronic low back pain with a neuropathic component: A randomized, controlled, open-label, phase 3b/4 trial. Pain Practice. 2015; doi: 10.1111/papr.12361
25. Bennett RM, Schein J, Kosinski MR, Hewitt DJ, Jordan DM, Rosenthal NR. Impact of fibromyalgia pain on health-related quality of life before and after treatment with tramadol/acetaminophen. Arthritis and Rheumatism. 2005; 53(4):519–527. [PubMed: 16082646]
26. Binsfeld H, Szczepanski L, Waechter S, Richarz U, Sabatowski R. a randomized study to demonstrate noninferiority of once-daily OROS[®] hydromorphone with twice-daily sustained-release oxycodone for moderate to severe chronic non-cancer pain. Pain Practice: The Official Journal of World Institute of Pain. 2010; 10(5):404–415. [PubMed: 20384968]

27. Gajria K, Kosinski M, Schein J, Kavanagh S, Dubois D. health-related quality-of-life outcomes in patients treated with push-pull OROS hydromorphone versus Extended-release oxycodone for chronic hip or knee osteoarthritis pain: A randomized, open-label, parallel-group, multicenter study. *The Patient*. 2008; 1(3):223–238. [PubMed: 22272928]
28. Gordon A, Callaghan D, Spink D, et al. Buprenorphine transdermal system in adults with chronic low back pain: A randomized, double-blind, placebo-controlled crossover study, followed by an open-label extension phase. *Clinical Therapeutics*. 2010; 32(5):844–860. [PubMed: 20685494]
29. Kalso E, Simpson KH, Slappendel R, Dejonckheere J, Richarz U. Predicting long-term response to strong opioids in patients with low back pain: Findings from a randomized, controlled trial of transdermal fentanyl and morphine. *BMC Medicine*. 2007; 5:39. [PubMed: 18154644]
30. Allan L, Richarz U, Simpson K, Slappendel R. Transdermal fentanyl versus sustained release oral morphine in strong-opioid naive patients with chronic low back pain. *Spine*. 2005; 30(22):2484–2490. [PubMed: 16284584]
31. Karlsson M, Berggren AC. Efficacy and safety of low-dose transdermal buprenorphine patches (5, 10, and 20 µg/h) versus prolonged-release tramadol tablets (75, 100, 150, and 200 mg) in patients with chronic osteoarthritis pain: A 12-week, randomized, open-label, controlled, parallel-group non-inferiority study. *Clinical Therapeutics*. 2009; 31(3):503–513. [PubMed: 19393841]
32. Khoromi S, Cui L, Nackers L, Max MB. Morphine, nortriptyline and their combination vs. placebo in patients with chronic lumbar root pain. *Pain*. 2007; 130(1–2):66–75. [PubMed: 17182183]
33. Ma K, Jiang W, Zhou Q, Du DP. The efficacy of oxycodone for management of acute pain episodes in chronic neck pain patients. *International Journal of Clinical Practice*. 2008; 62(2):241–247. [PubMed: 18070045]
34. Miller K, Yarlal A, Wen W, et al. Buprenorphine transdermal system and quality of life in opioid-experienced patients with chronic low back pain. *Expert Opinion on Pharmacotherapy*. 2013; 14(3):269–277. [PubMed: 23374027]
35. Nicholson B, Ross E, Sasaki J, Weil A. Randomized trial comparing polymer-coated extended-release morphine sulfate to controlled-release oxycodone HCl in moderate to severe nonmalignant pain. *Current Medical Research and Opinion*. 2006; 22(8):1503–1514. [PubMed: 16870075]
36. Pedersen L, Borchgrevink PC, Breivik HP, Fredheim OM. A randomized, double-blind, double-dummy comparison of short- and long-acting dihydrocodeine in chronic non-malignant pain. *Pain*. 2014; 155(5):881–888. [PubMed: 24345428]
37. Peloso PM, Fortin L, Beaulieu A, Kamin M, Rosenthal NR. Analgesic efficacy and safety of tramadol/acetaminophen combination tablets (Ultracet®) in treatment of chronic low back pain: A multicenter, outpatient, randomized, double blind, placebo controlled trial. *The Journal of Rheumatology*. 2004; 31(12):2454–2463. [PubMed: 15570651]
38. Rauck RL, Bookbinder SA, Bunker TR, et al. A randomized, open-label, multicenter trial comparing once-a-day AVINZA® (morphine sulfate extended-release capsules) versus twice-a-day OxyContin® (oxycodone hydrochloride controlled-release tablets) for the treatment of chronic, moderate to severe low back pain: Improved physical functioning in the ACTION trial. *Journal of Opioid Management*. 2007; 3(1):35–43. [PubMed: 17367093]
39. Ruoff GE, Rosenthal N, Jordan D, Karim R, Kamin M. Tramadol/acetaminophen combination tablets for the treatment of chronic lower back pain: A multicenter, randomized, double-blind, placebo-controlled outpatient study. *Clinical Therapeutics*. 2003; 25(4):1123–1141. [PubMed: 12809961]
40. Ueberall MA, Eberhardt A, Mueller-Schwefe GHH. Quality of life under oxycodone/naloxone, oxycodone, or morphine treatment for chronic low back pain in routine clinical practice. *International Journal of General Medicine*. 2016; 9:39–51. [PubMed: 26966387]
41. Watson CP, Moulin D, Watt-Watson J, Gordon A, Eisenhoffer J. Controlled-release oxycodone relieves neuropathic pain: A randomized controlled trial in painful diabetic neuropathy. *Pain*. 2003; 105(1–2):71–78. [PubMed: 14499422]
42. Yarlal A, Miller K, Wen W, et al. A randomized, placebo-controlled study of the impact of the 7-day buprenorphine transdermal system on health-related quality of life in opioid-naïve patients with moderate-to-severe chronic low back pain. *Journal of Pain*. 2013; 14(1):14–23. [PubMed: 23200931]

43. Hedges, LV., Olkin, I. Statistical methods for metaanalysis. San Deigo, CA: Academic Press; 1985.
44. Ahn SB, JB. Incorporating quality scores in meta-analysis. *Journal of Educational and Behavioral Statistics*. 2011; 36(5):555–585.
45. Wilson, D. Practical meta-analysis effect size calculator 2016. 2016. http://www.campbellcollaboration.org/resources/effect_size_input.php
46. EpiGear Internation Pty Ltd. Meta XL (5.0). 2015.
47. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials*. 1986; 7(3): 177–188. [PubMed: 3802833]
48. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003; 327(7414):557–560. [PubMed: 12958120]
49. Sawilowsky S. New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*. 2009; 8(2):597–599.
50. Kraemer HC, Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. *Biological Psychiatry*. 2006; 59(11):990–996. [PubMed: 16368078]
51. Cohen, J. Statistical power analysis for the behavioral sciences. New York: Academic Press; 1988.
52. Frich LM, Sorensen J, Jacobsen S, Fohlmann B, Hojsted J. Outcomes of follow-up visits to chronic nonmalignant pain patients. *Pain Management Nursing: Official Journal of The American Society of Pain Management Nurses*. 2012; 13(4):223–235. [PubMed: 23158704]
53. Sacks HS, Chalmers TC, Smith H. Randomized versus historical controls for clinical trials. *American Journal of Medicine*. 1982; 72:233–240. [PubMed: 7058834]
54. Schulz KF, Chalmers I, Hayes R, Altman DG. Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *Journal of the American Medical Association*. 1995; 273:408–412. [PubMed: 7823387]

Fig. 1.

Flow diagram for the selection of studies. The primary reason for study exclusion is only listed once even though some could have been excluded for multiple reasons

Fig. 2.

Forest plot from the random-effects model for Physical Component Summary (PCS) scores for all eligible studies. OA1: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs with two comparators each. OA1 is the first osteoarthritis study, where OA1.1 is tapentadol versus placebo and OA1.2 is long-acting oxycodone versus placebo. OA2: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs with two comparators each. OA2 is the second osteoarthritis study, where OA2.1 is tapentadol versus placebo and OA2.2 is long-acting oxycodone versus placebo. LBP: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs with two comparators each. LBP is the lower back pain study, where LBP1 is tapentadol versus placebo and LBP2 is long-acting oxycodone versus placebo. *ES* effect size

Fig. 3.
Funnel plot for Physical Component Summary (PCS) random-effects model in all eligible studies

Fig. 4.

Forest plot from the random-effects model for Mental Component Summary (MCS) scores for all eligible studies. OA1: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs with two comparators each. OA1 is the first osteoarthritis study, where OA1.1 is tapentadol versus placebo and OA1.2 is long-acting oxycodone versus placebo. OA2: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs with two comparators each. OA2 is the second osteoarthritis study, where OA2.1 is tapentadol versus placebo and OA2.2 is long-acting oxycodone versus placebo. LBP: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs with two comparators each. LBP is the lower back pain study, where LBP1 is tapentadol versus placebo and LBP2 is long-acting oxycodone versus placebo. *ES* effect size

Fig. 5.
Funnel plot for Mental Component Summary (MCS) random-effects model for all eligible studies

Fig. 6.

Forest plot from the random-effects model for Physical Component Summary (PCS) scores (opioid versus placebo studies). OA1: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs with two comparators each. OA1 is the first osteoarthritis study, where OA1.1 is tapentadol versus placebo and OA1.2 is long-acting oxycodone versus placebo. OA2: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs with two comparators each. OA2 is the second osteoarthritis study, where OA2.1 is tapentadol versus placebo and OA2.2 is long-acting oxycodone versus placebo. LBP: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs with two comparators each. LBP is the lower back pain study, where LBP1 is tapentadol versus placebo and LBP2 is long-acting oxycodone versus placebo. *ES* effect size

Fig. 7.

Forest plot from the random-effects model for Mental Component Summary (MCS) scores (opioid versus placebo studies). OA1: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs with two comparators each. OA1 is the first osteoarthritis study, where OA1.1 is tapentadol versus placebo and OA1.2 is long-acting oxycodone versus placebo. OA2: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs with two comparators each. OA2 is the second osteoarthritis study, where OA2.1 is tapentadol versus placebo and OA2.2 is long-acting oxycodone versus placebo. LBP: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs with two comparators each. LBP is the lower back pain study, where LBP1 is tapentadol versus placebo and LBP2 is long-acting oxycodone versus placebo. *ES* effect size

Table 1

Characteristics for all studies included in the systematic review

Author	Year	Country	Intervention and comparator	Participants	HRQoL measures
Afilalo [23] (OA1)	2013	Multisite—USA, Oceania	Opioid versus placebo: tapentadol versus placebo for 12 weeks Opioid versus placebo: oxycodone CR versus placebo for 12 weeks	Chronic osteoarthritis; tap ($n = 181$, age 58.4 y $\pm$ 10.9, 63% female); oxy ($n = 118$, age 58.2 y $\pm$ 10.3, 59% female); plac ($n = 203$, age 58.2 y $\pm$ 9.2, 59% female)	SF-36; PCS and MCS both reported
Afilalo [23] (OA2)	2013	Multisite—USA, Oceania	Opioid versus placebo: tapentadol versus placebo for 12 weeks Opioid versus placebo: oxycodone CR versus placebo for 12 weeks	Chronic osteoarthritis; tap ($n = 179$, age 62.4 y $\pm$ 9.4, 72% female); oxy ($n = 119$, age 61.8 y $\pm$ 9.1, 66% female); plac ($n = 215$, age 62.2 y $\pm$ 9.4, 76% female)	SF-36; PCS and MCS both reported
Afilalo [23] (LBP)	2013	Multisite—USA, Oceania	Opioid versus placebo: tapentadol versus placebo for 12 weeks Opioid versus placebo: oxycodone CR versus placebo for 12 weeks	Chronic low back pain; tap ($n = 166$, age 49.4 y $\pm$ 13.2, 61% female); oxy ($n = 133$, age 50 y $\pm$ 14.2, 55% female); plac ($n = 152$, 50.4 y $\pm$ 14.1, 58% female)	SF-36; PCS and MCS both reported
Baron [24]	2015	Multisite—Europe	Opioid versus opioid: tapentadol versus oxycodone/naloxone for 12 weeks	Chronic low back pain; tap ($n = 130$, age 58.1 y $\pm$ 11.5, 59% female); oxy ($n = 125$, age 58.4 y $\pm$ 12.2, 66% female)	SF-12; PCS and MCS both reported
Bennett [25]	2005	Multisite—USA	Opioid versus placebo: tramadol/APAP versus placebo for 13 weeks	Fibromyalgia; tram/APAP ($n = 156$, age 49 y $\pm$ 11, 93% female); plac ($n = 157$, age 51 y $\pm$ 10, 95% female)	SF-36; PCS not reported as “significantly difference” and MCS reported as “no significant differences”
Binsfeld [26]	2010	Multisite—Europe	Opioid versus Opioid: hydromorphone ER versus oxycodone SR scheduled for 52 weeks	Non-specific pain; Hydro ($n = 254$, age 57.1 y $\pm$ 13.1, 56% female); oxy ($n = 250$, 57.5 y $\pm$ 12.8, 60% female)	SF-36; PCS reported; MCS not reported
Gajria [27]	2008	Multisite—Europe	Opioid versus opioid: hydromorphone ER versus oxycodone SR for 6 weeks	Chronic osteoarthritis Hydro ($n = 64$, age 62.9 y $\pm$ 10.3, 77% female); oxy ($n = 60$, 64.2 y $\pm$ 13.1, 68% female)	WOMAC Global and Physical Component reported
Gordon [28]	2010	Canada	Opioid versus placebo: buprenorphine transdermal versus placebo for 8 weeks	Chronic low back pain; both ($n = 52$, age 50.7 y $\pm$ 11.9, 62% female)	SF-36; PCS and MCS not reported—“no significant differences”
Kalso [29, 30] ^a	2005	Multisite—Europe	Opioid versus opioid: fentanyl TD versus morphine SR for 56 weeks	Chronic low back pain; fent ($n = 338$, age 53.4 y, 61% female); mor ($n = 342$, 54.7 y, 62% female)	SF-36; PCS and MCS reported
Karlsson [31]	2009	Multisite—Europe	Opioid versus opioid: buprenorphine transdermal system (BTDS) vs tramadol LA for 12 weeks	Chronic osteoarthritis; BTDS ($n = 69$, age 64.4 y $\pm$ 11.1, 59% female); tram ($n = 65$, 64.2 y $\pm$ 9.3, 54% female)	EQ-5D; reported “did not differ significantly between groups.”
Khoromi [32]	2007	USA	Opioid versus placebo: morphine SR versus buprenorphine (active, non-opioid placebo) for 9 weeks	Chronic leg pain; mor ($n = 28$, age 53 y, 50% female); benz ($n = 28$, 52.5 y, 50% female)	SF-36; PCS and MCS reported
Ma [33]	2008	China	Opioid versus placebo: oxycodone CR versus placebo for 4 weeks	Chronic neck pain; oxy ($n = 58$, age 58.2 y $\pm$ 12.6, 31% female); plac ($n = 58$, 53.1 y $\pm$ 16.4, 45% female)	SF-36; PCS and MCS reported

Author	Year	Country	Intervention and comparator	Participants	HRQoL measures
Miller [34]	2013	Multisite—USA	Opioid versus opioid: buprenorphine BTDS high dose versus BTDS low dose for 12 weeks Opioid versus opioid: BTDS high dose versus oxycodone IR for 12 weeks	Chronic low back pain; high ($n = 143$, age 50.4 y ± 11.9 , 52% female); low ($n = 131$, 50.2 y ± 12.9 , 46% female); oxy ($n = 163$, 49.5 y ± 12.4 , 45% female)	SF-36; PCS reported; MCS reported as “no significant differences”
Nicholson [35]	2006	Multisite—USA	Opioid versus opioid: morphine ER versus oxycodone CR for 24 weeks	Non-specific pain; mor ($n = 35$, age 51.3 y, 63% female); oxy ($n = 38$, 51.3 y, 41% female)	SF-36v2; PCS and MCS both reported as “no significant differences”
Pedersen [36]	2014	Norway	Opioid versus opioid: codeine SR versus codeine IR for 8 weeks	Non-specific pain; SR ($n = 28$, age 49 y, 61% female); IR ($n = 30$, 49 y, 47% female)	SF-8; PCS and MCS reported
Peloso [37]	2004	Multisite—North America	Opioid versus placebo: tramadol versus placebo for 13 weeks	Chronic low back pain; tram ($n = 163$, age 57.5 y ± 11.5 , 64% female); plac ($n = 163$, 57.5 y ± 13.6 , 61% female)	SF-36; PCS and MCS reported
Rauck [38]	2007	USA	Opioid versus opioid: morphine ER versus oxycodone CR for 16 weeks	Chronic low back pain; mor ($n = 43$, age and % female not specified); oxy ($n = 54$, age and % female not specified)	SF-12; PCS and MCS both reported as “no significant differences”
Ruoff [39]	2003	Multisite—USA	Opioid versus placebo: tramadol/APAP versus placebo scheduled for 13 weeks	Chronic low back pain; tram ($n = 157$, 54.1 y ± 12 , 59% female)	SF-36; PCS and MCS reported
Ueberall [40]	2016	Germany	Opioid versus opioid: morphine IR versus oxycodone/naloxone for 12 weeks Opioid versus opioid: oxycodone IR versus oxycodone/naloxone for 12 weeks	Chronic low back pain; mor ($n = 300$, age 46.5 y ± 9.3 , 56% female); oxy ($n = 300$, 46.7 y ± 9.9 , 55% female); oxy/nal ($n = 301$, 46.1 y ± 9.9 , 56% female)	SF-12; PCS reported as having “significantly greater achievements” and MCS reported as “significantly better”
Watson [41]	2003	Canada	Opioid versus placebo: oxycodone CR versus benzotropine (active, non-opioid placebo) cross-over scheduled for 4 weeks	Diabetic-related nerve pain; both ($n = 36$, age 63 y ± 9.4 , 47% female)	SF-36; PCS and MCS reported
Yarlas [42]	2013	Multisite	Opioid versus placebo: buprenorphine BTDS versus placebo for 12 weeks	Chronic low back pain; BTDS ($n = 170$, age and % female not specified); plac ($n = 199$, age and % female not specified)	SF-36v2; PCS was significant in bar chart with no values reported and MCS was not significant in bar chart with no values reported

Inter intervention, *Comp* comparator, *Place* placebo, *Mor* morphine, *NaI* naloxone, *SF-36* 36-item Short Form Health Survey, *PCS* Physical Component Summary, *MCS* Mental Component Summary, *ER* extended release, *SR* sustained release, *CR* controlled release, *WOMAC* Western Ontario and McMaster Universities Arthritis Index is a specific HRQoL instrument, *IT* intrathecal, *LD* low dose, *TD* transdermal, *APAP* acetaminophen

OA1: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs. The first study is osteoarthritis study #1

OA2: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs. The second study is osteoarthritis study #2

LBP: Afilalo 2013 is one manuscript which presents HRQoL data from three RCTs. The third study is lower back pain

^aThe citation Kalso 2005 references Allan 2005 which was used for partial data abstraction despite not being included via the search strategy

Table 2

Cochrane risk of bias tool scores for each study included in the systematic review

Author	Year	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Overall risk of bias
Afilalo [23]	2013	Low	Low	Unclear	Unclear	Low	High	High
Baron [24]	2015	Unclear	High	High	Unclear	Low	High	High
Bennett [25]	2005	Unclear	Low	Low	Low	Low	Low	Low
Binsfeld [26]	2010	Unclear	High	High	Unclear	Low	High	High
Gajria [27]	2008	Low	High	High	Unclear	High	Low	High
Gordon [28]	2010	Low	Low	Low	Low	Low	Low	Low
Kalso [29, 30] *	2005	Low	Unclear	High	High	Low	Low	High
Karlsson [31]	2009	Low	High	High	High	Low	High	High
Khoromi [32]	2007	Low	Low	Low	Unclear	Low	Low	Low
Ma [33]	2008	Unclear	Unclear	Unclear	Unclear	High	Low	High
Miller [34]	2013	Unclear	Low	Low	Unclear	High	Unclear	Low
Nicholson [35]	2006	Low	High	High	Unclear	Low	High	High
Pedersen [36]	2014	Low	Low	Low	Low	Low	Low	Low
Peloso [37]	2004	Low	Low	Low	Low	Low	Low	Low
Rauk [38]	2007	Low	High	High	High	High	High	High
Ruoff [39]	2003	Low	Low	Low	Low	Low	Low	Low
Ueberall [40]	2016	Low	High	High	Low	High	High	High
Watson [41]	2003	Low	Low	Low	Low	High	Low	High
Yarlas [42]	2013	Unclear	Low	High	Unclear	Low	Low	High

Table 3

Changes in health-related quality of life outcomes

Analysis	Studies	Comparisons	<i>g</i> (95% CI)	<i>Q</i> (<i>p</i>)	<i>I</i> ² (95% CI)
PCS scores for all eligible studies	8	13	0.22 (0.11, 0.32)	30.86 (0.002) *	61.1 (28.8, 78.8)
MCS scores for all eligible studies	6	11	-0.17 (-0.52, 0.17)	247.2 (<0.001) *	96.0 (94.3, 97.2)
PCS scores for opioid versus placebo studies only	3	8	0.18 (0.08, 0.28)	12.0 (0.099) *	41.8 (0.0, 74.3)
MCS scores for opioid versus placebo studies only	3	8	-0.05 (-0.18, 0.08)	19.1 (0.007) *	63.3 (21.2, 82.9)

Boldfaced items indicate statistical significance with non-overlapping 95% confidence intervals

PCS Physical Component Summary, MCS Mental Component Summary, *g* Hedge's *g* effect size, *CI* 95% confidence interval, *Q* Cochran's *Q*, *p* *p*-value, *I*² inconsistency (%), and 95% confidence interval

* Statistically significant (*p* 0.10)