

Efficacy and harms of long-term opioid therapy in chronic non-cancer pain: Systematic review and meta-analysis of open-label extension trials with a study duration ≥ 26 weeks

Patric Bialas¹ | Christoph Maier² | Petra Klose³ | Winfried Häuser^{4,5}

¹Department of Anesthesiology, Universitätskliniken des Saarlandes, Homburg/Saar, Germany

²Department of Pain Medicine, Ruhr-Universität Bochum, Bochum, Germany

³Department Internal and Integrative Medicine, Kliniken Essen-Mitte, Faculty of Medicine, University of Duisburg-Essen, Essen, Germany

⁴Health Care Center for Pain Medicine and Mental Health, Saarbrücken, Germany

⁵Department Psychosomatic Medicine and Psychotherapy, Technische Universität München, München, Germany

Correspondence

Winfried Häuser, Health Care Center for Pain Medicine and Mental Health, D-66119 Saarbrücken, Germany.
Email: w.haeuser@web.de

Abstract

Background and Objective: This updated systematic review evaluated the efficacy, acceptability and safety of long-term opioid therapy (LTOT) for chronic non-cancer pain (CNCP).

Databases and Data Treatment: Clinicaltrials.gov, CENTRAL and MEDLINE until June 2019. We included open-label extension trials with a study duration ≥ 26 weeks of RCTs with ≥ 2 weeks duration. Pooled estimates of event rates of categorical data and standardized mean differences (SMD) of continuous variables were calculated using a random effects model.

Results: We added four new studies with 1,154 participants for a total of 15 studies with 3,590 participants. Study duration ranged between 26 and 156 weeks. Studies included patients with low back, osteoarthritis and neuropathic pain. The quality of evidence for every outcome was very low. 31.1% (95% Confidence interval [CI] 23.0%–40.7%) of patients randomized at baseline finished the open label period. 14.1% (95% CI 10.9%–19.4%) of patients dropped out to due adverse events. In 6.3% (95% CI 3.9%–10.1%) of patients serious adverse events and in 2.7% (95% CI 1.5%–4.7%) aberrant drug behaviour were noted. 0.5% (95% CI 0.2%–1.4%) of patients died.

Conclusions: Within the context of open-label extension studies, opioids maintain reduction of pain and disability and are rather well tolerated and safe. LTOT can be considered in carefully selected and monitored patients with low back, osteoarthritis and neuropathic pain who experience a clinically meaningful pain reduction with at least tolerable adverse events in short-term opioid therapy.

Significance: There is very low quality evidence of the long-term efficacy, tolerability and safety of opioids for chronic low back, osteoarthritis and diabetic polyneuropathic pain within the context of open-label extension studies of randomized controlled trials. Drop out rate due to adverse events and deaths increase with study duration. One-third of patients profit from LTOT. Long-term opioid therapy can be considered in some carefully selected and monitored patients.

1 | INTRODUCTION

The rates of long-term (>3 months; von Korff et al., 2008) opioid therapy (LTOT) in chronic non-cancer pain (CNCP) are increasing in western countries including Germany (Häuser, Schug, & Furlan, 2017). The increase of opioid prescriptions was associated with an increase of aberrant drug behaviour (abuse, diversion, addiction) and deaths (overdose, accident) in North America, the so-called opioid epidemic (Volkow, Jones, Einstein, & Wargo, 2019). A systematic review found that the rates of misuse averaged between 21% and 29% and of addiction averaged between 8% and 12% in patients of US pain clinics (Vowles et al., 2015). Two US and one French claims database studies (Chenaf et al., 2019; Ray, Chung, Murray, Hall, & Stein, 2016; Zeng et al., 2019) found an increased mortality risk associated with LTOT.

In the view of the North American opioid crisis, recent US systematic reviews on opioids for CNCP applied standards for the proof of the efficacy and safety of opioids which has not been required for any pain drug before. A recent US review stated that conclusions on the effectiveness of LTOT for chronic pain are not possible due to the paucity of research to date. The authors did not find an RCT >3 months comparing opioids to placebo or an active comparator. In addition, they stated that the studies were not designed to assess serious harms of opioid therapy such as abuse and diversion (Chou et al., 2015).

The lack of long-term RCTs of opioids might be due to the fact that drug agencies such as the European Medicines Agency (EMA) require for the proof of efficacy RCTs of at least 12 weeks double blind duration (European Medicines Agency, 2017). These requirements are met for opioids for chronic low back pain (Petzke et al., 2015), neuropathic pain (Sommer et al., 2015) and osteoarthritis pain (Schaefer et al., 2015). EMA requires, that long-term efficacy and safety should be tested in uncontrolled long-term trials, for instance in an open label extension phase during 6 to 12 months without placebo (European Medicines Agency, 2017). These databases have been neglected in the recent US and Canadian evidence-based guidelines on opioid therapy for chronic pain (Busse et al., 2017; Dowell, Haegerich, & Chou, 2016).

For the second update on the German guidelines on LTOT for CNCP (LONTS; Häuser, Bock, et al., 2015) we updated our review of open label extension studies of randomized controlled trials of opioids with a study duration of ≥ 2 weeks for any type of CNCP with patients with any age (Häuser, Bernardy, & Maier, 2015). Specifically we studied, how many patients remained on opioid therapy, reported a sustained reduction of pain and disability and experienced serious harms (serious adverse events, death, aberrant drug behaviour) in the long term (≥ 6 months).

2 | MATERIALS AND METHODS

The review was performed according to the PRISMA statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; Moher et al., 2009).

2.1 | Study protocol

Methods of analysis and inclusion criteria were specified in advance (PROSPERO CRD42019124303).

2.1.1 | Criteria for considering studies for this review

2.1.1.1 Types of participants

Patients of any age with pain due to any cause other than cancer lasting for at least 3 months prior to trial enrolment.

Types of interventions

We considered trials with the following opioids and their administration compared to placebo: a. Opioids given by oral, buccal and transdermal routes. b. opioids administered as abuse deterrent formulations, for example, in combination with naloxone. c. Tramadol, a centrally acting, synthetic opioid analgesic with two complementary mechanisms of action: binding of parent and M1 metabolite to μ -opioid receptors and inhibition of reuptake of norepinephrine and serotonin. d. Tapentadol, a drug with two mechanisms of action: μ -receptor agonist and norepinephrine reuptake inhibitor. The reason for including both latter drugs into this review was that they are classified as opioids by German medicine agencies.

We excluded trials a. that examined opioids given by an intravenous route due to the invasive nature of the therapy and its limited clinical relevance in the outpatient setting. We did not assess the effectiveness of opioids delivered by neuraxial implantable pumps, as this has been discussed elsewhere (Noble et al., 2010). b. in which analgesics other than opioid agonists were combined with opioids (e.g. tramadol with acetaminophen), because it is not possible to disentangle the effects of the opioids from those of the other analgesic. If only used as rescue analgesic the combination was allowed. c. in which a defined opioid was compared to the same opioid with ADFs (e.g. oxycodone with and without naloxone) or in which two opioids combined were compared to a single opioid without a placebo group. d. with opioid receptor agonist/N-methyl-D-aspartate (NMDA) antagonists (e.g. levorphanol) because these drugs are not available in Germany. e. with methadone and levomethadone because these drugs are nearly primarily used to treat opiate addiction in Germany. f. with drugs under

development (such as cepranopadol) which have not been approved by the EMA.

2.1.1.3. Types of studies

We included open label extension studies of RCTs. Study duration of the RCTs should have been at least 2 weeks and of the open label extension phase at least 6 months. Open label duration should include at least 20 patients. Studies should have reported at least one of the outcomes defined below. We excluded registered open label studies without prior double blind randomized period and pre–post case-series studies because these studies are associated with a higher risk of bias than open label extension studies of RCTs (Moore et al., 2010). We excluded studies with cancer and non-cancer pain samples if the results for non-cancer pain were not reported separately.

Types of outcome measures

Efficacy. Changes of pain intensity at the end of open label compared to the end of randomized trial of patients on study medication at the end of randomized period.

Changes of disability at the end of open label compared to the end of randomized trial of patients on study medication at the end of randomized period.

Number of patients who were on opioids at the end of open label (related to the number of patients randomized at baseline and to the number of patients included into open label).

Number of patients which dropped out due to lack of efficacy.

Tolerability. Number of patients which dropped out due to adverse events.

Safety. Number of serious adverse events (SAE).

Number of patients with aberrant drug behaviour.

Number of deaths.

Outcome measures must have been validated or used as a standard of care to be included in the analyses. In addition to these general inclusion criteria, we employed two criteria for efficacy outcomes: (a) pain and disability outcomes must have been patient-reported; (b) outcome data must not have been collected retrospectively (for example, post-treatment surveys/questionnaires).

2.1.2 | Electronic searches

The search included CENTRAL, Medline and clinicaltrials.gov from inception to 16 June 2019. Our search included all languages. We searched all databases with the search terms “open-label extension” AND (“buprenorphine” OR

“codeine” OR “fentanyl” OR “hydrocodone” OR “hydromorphone” OR “morphine” OR “oxycodone” OR “oxymorphone” OR “tapentadol” OR “tilidine” or “tramadol”).

2.2 | Measures of treatment effect

Standardized mean differences (SMD) of continuous variables were calculated using means and standard deviations for each intervention. Pooled estimates of event rates of categorical data (e.g. drop out due to SAE) were calculated using a random effects model. Confidence intervals (95% CI) were calculated for all summary data.

We used the I^2 statistic to identify heterogeneity. Combined results with $I^2 > 50\%$ were considered substantially heterogeneous (Higgins, Churchill, Chandler, & Cumpston, 2017).

2.3 | Data collection and analysis

2.3.1 | Selection of studies

Two authors (BP, WH) selected the studies. Disagreements on study selection were resolved by consensus. If needed, a third review author was involved (CM).

2.3.2 | Data extraction and management

Two pairs of review authors extracted the data from the full text articles and entered the data independently in standard extraction forms (KB, WH; PB, WH). We extracted characteristics of patients and studies, description of the experimental and control, co-interventions, affiliations of the authors and sponsoring of the study. Disagreements were resolved by consensus. If needed, a third review author was involved (CM).

2.3.3 | Assessment of risk of bias in included studies

Two pairs of authors (KB, WH; PB, WH) independently assessed the risk of bias in each trial assessed using six domains recommended by the Cochrane Collaboration: selection bias, performance bias, detection bias, attrition bias, reporting bias (Higgins et al., 2017). We slightly modified one item of the tool (selection bias) to adapt to the setting of an open label extension trial (see Table S1). The criteria were scored as “yes”, “no” or “unclear”. Any disagreements were resolved by discussion. If needed, a third review author was involved (CM). We defined a high quality study that fulfilled five to six, a moderate quality study that fulfilled four to three and a low quality study that fulfilled zero to two of the six validity criteria.

2.3.4 | Grading of evidence

We used GRADE (Langendam et al., 2013) to assess the overall quality of evidence. The quality of evidence was downgraded by one level for each of the following factors that were encountered:

- Limitations of study design: >50% of the participants of studies with a high risk of bias
- Inconsistency of results: $I^2 > 50\%$
- Indirectness: We assessed whether the question being addressed in this systematic review was different from the available evidence regarding the population in routine clinical care if patients with clinically relevant internal diseases [heart, lung, kidney, liver] and/or major mental disorders [history of substance abuse or major depression] were excluded in >50% of participants
- Imprecision: There was only one trial or when there was more than one trial, the total number was <400 patients or when the pooled estimate of effect included no effect.

We categorized the quality of evidence as follows:

- High (++++): we are very confident that the true effect lies close to that of the estimate of the effect.
- Moderate (+++): we are moderately confident in the effect estimate; the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
- Low (++) : our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect.
- Very low (+): we have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of effect; any estimate of effect is very uncertain.

2.4 | Subgroup analysis

Provided that at least two studies were available, subgroup analyses were predefined for type of opioid and the type of chronic pain syndrome for the outcomes pain intensity, drop out due to adverse events and serious adverse events. These subgroup analyses were also used to examine potential sources of clinical heterogeneity.

2.5 | Sensitivity analyses

Sensitivity analyses were predefined by excluding studies with imputed means and SDs.

2.6 | Assessment of publication bias

For analyses with at least 10 studies, we used the Egger intercept test (Egger, Smith, Schneider, & Minder, 1997) and the Begg rank correlation test for funnel plot asymmetry (Begg & Mazumdar, 1994) at the significance level $p < .05$. In addition, publication bias was controlled for by computing "safe-n rates" (Orwin, 1983).

2.7 | Metaregression

The impact study duration on outcomes of efficacy, tolerability and safety was tested by metaregression, which was performed using the random-effects model. τ^2 variance was calculated by the method of maximum likelihood. Goodness of fit (test that unexplained variance is zero) was calculated for the model (Comprehensive meta-analysis, 2010).

2.8 | Software

Comprehensive meta-analysis (Biostat) model (Comprehensive meta-analysis, 2010 and RevMan Analysis (RevMan 5.3) of the Cochrane Collaboration software (Review Manager, 2014) were used for statistical analyses.

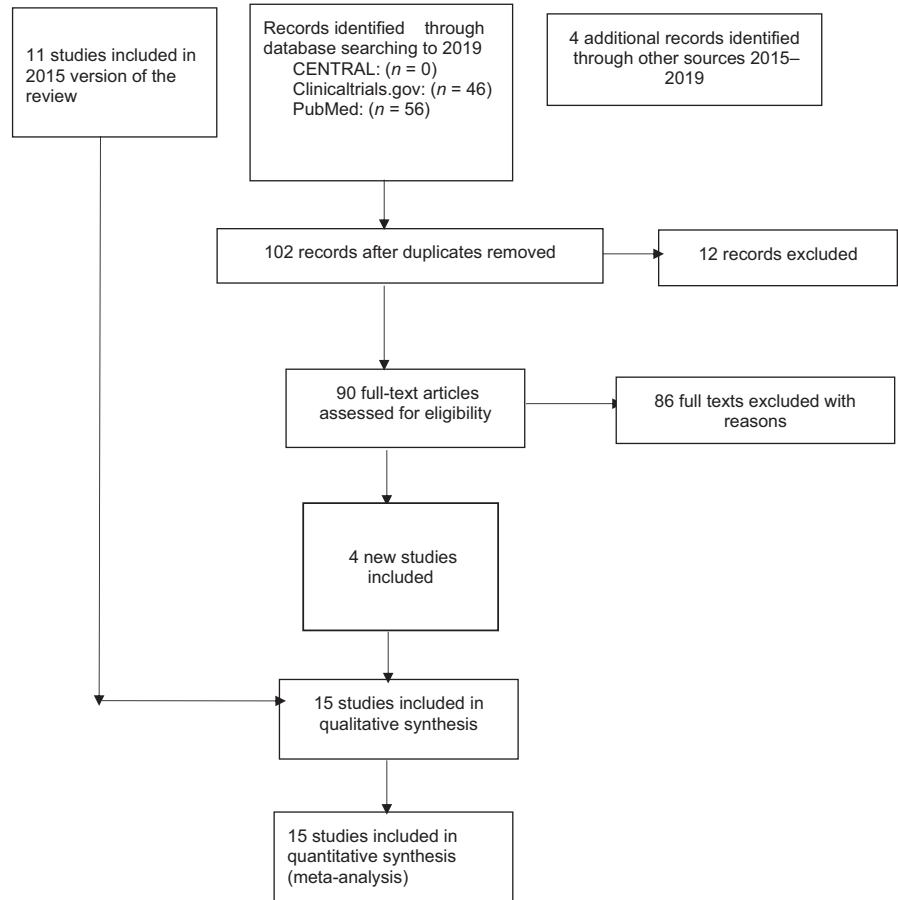
3 | RESULTS

3.1 | Search

The total number of included studies in the 2015 review was 11 (Caldwell et al., 2002; Cloutier et al., 2013; Gordon et al., 2010; Harati et al., 2000; Johnson & Johnson, 2010; McIlwain & Ahdieh 2005; Portenoy et al., 2007; Richarz, Waechter, Sabatowski, Szczepanski, & Binsfeld, 2013; Roth et al., 2000; Sandner-Kiesling et al., 2010; Thorne et al., 2008). One study which was only available in a database (Johnson & Johnson, 2010) had been published as full paper (Buynak et al., 2015). The updated searches (last performed June 16, 2019) produced 90 hits after duplicates were removed. We included four new studies with 1,154 participants (Blagden, Hafer, Duerr, Hopp, & Bosse, 2014; Hale, Zimmerman, Ma, and Malamut 2015; Hale, Urdaneta, Kirby, Xiang, & Rauck, 2017; Kawamata et al., 2019) and a total of 15 studies with 3,590 participants into the qualitative and quantitative analysis (see Figure 1).

3.2 | Included studies

The main characteristics of the studies are summarized in Table 1, for details see Results Table S1.

FIGURE 1 PRISMA flow diagram

3.2.1 | Settings

Six studies were conducted in USA, four studies in Canada and two studies each in Europe and in other continents and one study in Japan.

The open label duration of eight studies lasted 26 weeks, of five studies 52 weeks and of one study each 78 and 156 weeks.

3.2.2 | Types of opioids

Two studies tested buprenorphine one by transdermal and one by buccal route (5–40 µg/hr; reported mean dosage 14 µg/hr). One study tested hydrocodone (30 to 90 mg/day; the most frequent dosage in opioid naïve patients was 30 mg/day and in opioid experienced patients 90 mg/day). Two studies tested hydromorphone (8 to 32 mg/day; mean dosage in one study 17 mg/day). One study tested morphine (maximum dosage 90 mg/day, half of the patients used 60 mg/day). Six studies tested oxycodone between 20 and 140 mg/day (mean dosage in one study was 44 mg/day). One study tested oxycodone/naloxone (maximum dosage 120/60 mg/day; the mean dosage reported in one study was 44/22 mg/day). One study with oxymorphone (40 to 80 mg/day) reported an average dosage of 62 mg/day. One study

tested tapentadol (100 to 500 mg/day; mean dosage 368 mg/day). Two studies tested tramadol (maximum dosage 400 mg/day; no mean dosages reported).

Two studies included two opioids: Buynak et al. (2015) tested oxycodone and tapentadol). Richarz et al. (2013) tested hydromorphone and oxycodone. The remaining studies had two study arms (opioid vs. control).

All studies with on oral opioid used an extended release form.

3.2.3 | Types of CNCP

One study did not specify the type of CNCP. Four studies each included patients with osteoarthritis and low back pain. Three studies included both pain syndromes. Two studies included patients with diabetic polyneuropathy, osteoarthritis and low back pain. One study included patients with diabetic polyneuropathy.

3.2.4 | Participants

All studies included only adults. If reported, the mean age of the participants ranged between 52 and 62 years. If reported, the

TABLE 1 Overview of open label extension studies of randomized controlled trials included into the systematic review (grouped by type of opioid in alphabetical order)

| Reference Year Countries of study centres | Diagnosis | Number of patients randomized at study entry/ included into open-label/ finished open-label | Study medication Mean dosage (mean \pm SD) | Duration of open label trial |
|--|---|---|---|------------------------------|
| Buprenorphine | | | | |
| Gordon et al. (2010) USA, Canada | Low back pain | 79/42/26 Long-term completer: 32.9% | 7-day buprenorphine flexible 5 or 10 or 20 μ g/hr transdermal Mean dosage 14.3 \pm 5.7 μ g/hr | 26 weeks |
| Hale et al. (2017) USA | Low back pain | 506/435/158 Long-term completer: 32.9% | 7-day buprenorphine flexible 5 or 10 or 20 μ g/hr transdermal 52 (12.0%) reached an optimal dose of 300 μ g, 45 (10.3%) reached an optimal dose of 450 μ g, 141 (32.4%) reached an optimal dose of 600 μ g, 62 (14.3%) reached an optimal dose of 750 μ g, and 135 (31.0%) reached an optimal of 900 μ g | 48 weeks |
| Hydrocodone | | | | |
| Hale, Zimmerman, Ma, and Malamut (2015) USA | Low back pain | 329/183/126 Long-term completer: 38.3% | Hydrocodone extended release 30 or 45 or 60 or 90 mg/day Most frequent dosage 30 mg/day in opioid naive and 90 mg/day in opioid experienced patients | 26 weeks |
| Hydromorphone | | | | |
| Richarz et al. (2013) USA | Low back pain, musculoskeletal pain, neuropathic pain | 504/112/97* Long-term completer: 19.2%* *No separate data for both groups reported | Hydromorphone extended release flexible 8–32 mg/day oral or Oxycodone controlled-release flexible 20–80 mg/day oral Mean dosage hydromorphone 17.1 mg/day and oxycodone 44.6 mg/day | 28 weeks |
| Morphine | | | | |
| Caldwell et al. (2002) USA | Osteoarthritis pain | 295/181/86 Long-term completer: 29.2% | Morphine 30 mg/day extended release once daily in the morning or evening oral 42 (49%) remained on 30 mg/day morphine dose; 7 patients increased to 120 mg/day | 26 weeks |
| Oxycodone | | | | |
| Blagden et al. (2014) Europe and USA | Chronic non-cancer pain (no details provided) | 587/474/399 Long-term completer: 68.0% | Oxycodone and naloxone prolonged release flexible up to 120 mg/day Mean dosage not reported | 52 weeks |
| Cloutier et al. (2013) Canada | Low back pain | 83/50/40 Long-term completer: 49.2% | Oxycodone Naloxone controlled release oral flexible 20/10 mg/day or 30/15 mg/day or 40/20 mg/day Mean dosage at the end of the open label: 35.1/17.6 mg/day | 26 weeks |
| Kawamata et al. (2019) Japan | Low back pain | 83/75/54 Long-term completer: 65.1% | Oxycodone extended release flexible. Mean dosage during open label: 24.86 mg | 52 weeks |

(Continues)

TABLE 1 (Continued)

| Reference Year Countries of study centres | Diagnosis | Number of patients randomized at study entry/ included into open-label/ finished open-label | Study medication Mean dosage (mean ± SD) | Duration of open label trial |
|--|---|---|--|------------------------------|
| Portenoy et al. (2007) USA | Osteoarthritis and low back pain | NR/227/39 Long-term completer: Not calculable | Oxycodone controlled release flexible 20–140 mg/day Overall mean (±SD) daily dose s 52.5 (±38.5) mg. range: 10.0 to 293.5 mg/day | 156 weeks |
| Richarz et al. (2013) USA | Low back pain, musculoskeletal pain, neuropathic pain | 504/112/97* Long-term completer: 19.2%* *No separate data for both groups reported | Hydromorphone extended release flexible 8–32 mg/day or Oxycodone controlled-release flexible 20–80 mg/day oral Mean dosage hydromorphone 17.1 mg/day and oxycodone 44.6 mg/day | 28 weeks |
| Roth et al. (2000) USA | Osteoarthritis pain | 133/106/15 Long-term completer: 11.3% | Oxycodone controlled release flexible 20–80 mg/day The dose became constant at approximately 40 mg/day by week 16 | 26 weeks |
| Sandner Kiesling (2010) Europe | Osteoarthritis and low back pain | 463/258/243 Long-term completer: 52.5% | Oxycodone and naloxone prolonged release fixed 20 or 40 mg/day 35.6 ± 16.53 mg after 2 weeks to 43.7 ± 22.53 mg at the end of the extension phase | 52 weeks |
| Oxymorphone | | | | |
| McIlwain and Ahdieh (2005) USA | Osteoarthritis pain | 269/153/61 Long-term completer: 22.7% | Oxymorphone extended release flexible oral 40 to 80 mg/day (median average dose 62 mg/day at week 52) | 52 weeks |
| Tapentadol | | | | |
| Buynak et al. (2015) Australia, Europe, New Zealand and USA | Low back pain or osteoarthritis knee or hip pain | 1514/1154/698 Long-term completer: 46.1% | Tapentadol extended release flexible 200–500 mg/day oral The mean total daily dose of tapentadol was 368.2 mg/day | 52 weeks |
| Tramadol | | | | |
| Harati et al. (2000) USA | Painful Diabetic polyneuropathy | 1317/117/58 Long-term completer: 4.4% | Tramadol flexible 100–400 mg/day oral not reported | 26 weeks |
| Thorne et al. (2008) Canada | Osteoarthritis | 100/53/29 Long-term completer: 29.0% | Tramadol flexible 100–400 mg/d oral The mean final dose of tramadol was 313.2 ± 100.1 mg/day. compared with 330.2 ± 93.7 mg/day during the last week of double-blind active treatment | 26 weeks |

percentage of Caucasians ranged between 80% and 100% except one study which included only Asian patients. The percentage of women in the studies ranged between 44% and 76%. The number of patients included ranged between 42 and 1,174.

3.2.5 | Exclusion of clinically relevant internal diseases or mental disorders

Thirteen studies excluded patients with relevant internal diseases. Twelve studies excluded patients with a history of or current substance abuse.

3.2.6 | Funding and conflicts of interest

Thirteen studies reported sponsoring by pharmaceutical companies. No study received public funding. Ten author groups declared their conflicts of interest.

3.3 | Risk of bias in included studies

Risk of bias could not be properly assessed in all studies due to poor method reporting. In general, the risk of bias of all included studies was high for selection, performance and detection bias due to the study design. All studies were funded by the manufacturers of the drug (see Figure 2). Detailed information regarding risk of bias assessments of every study are given in Results Table S2.

3.4 | Effects of intervention

The quality of evidence was very low due to limitations of study design, indirectness, inconsistency (except the outcome aberrant drug behaviour) and high probability of a publication bias.

(Results are reported with 95% CI):

3.4.1 | Mean pain intensity at end of open-label versus end of double-blind period

Nine studies with 2,689 participants were entered into an analysis of mean pain intensity at the end of open-label versus at the end of double blind period of patients on study medication at the end of randomized period. Change in the pain intensity between these two study periods was SMD 0.06 [-0.03, 0.15]; ($p = .22$; $I^2 = 47$; see Results S1: figure 1). The association of study duration and the SMD was $\beta = 0.001$, $p = .24$; $\tau^2 = 0.01$, $p = .02$.

| | Selection bias | Allocation concealment (selection bias) | Blinding of participants and personnel (performance bias) | Blinding of outcome assessment (detection bias) | Incomplete outcome data (attrition bias) | Selective reporting (reporting bias) |
|-----------------------|----------------|---|---|---|--|--------------------------------------|
| Bladgen 2014 | ? | — | — | — | ? | ? |
| Buynak 2015 | ? | — | — | — | ? | ? |
| Caldwell 2002 | ? | — | — | — | ? | — |
| Cloutier 2013 | ? | — | — | — | ? | — |
| Gordon 2010 | ? | — | — | — | ? | — |
| Hale 2016 | ? | — | — | — | ? | — |
| Hale 2017 | ? | — | — | ? | ? | ? |
| Harati 2000 | ? | — | — | — | ? | ? |
| Kawamata 2019 | ? | — | — | — | ? | — |
| McIlwain 2005 | ? | — | — | — | ? | — |
| Portenoy 2007 | ? | — | — | — | ? | — |
| Richarz 2013 | ? | — | — | — | ? | — |
| Roth 2000 | ? | — | — | — | — | ? |
| Sandner-Kiesling 2010 | ? | — | — | — | ? | ? |
| Thorne 2008 | ? | — | — | — | — | — |

FIGURE 2 Risk of bias summary

3.4.2 | Disability

Three studies with 1,066 participants were entered into an analysis of mean disability at the end of open-label versus at the end of double-blind period of patients on study medication at the end of randomized period. Change in physical function between these two study periods was SMD -0.12 [-0.24 , 0.00] ($p = .05$; $I^2 = 0$; see Results S1: figure 2). With only three studies available we did not perform regression analysis.

3.4.3 | Patients entering/finishing open label period

Fifteen studies with 3,590 participants were entered into an analysis of patients entering/finishing open label period. 56.8% (45.0% to 67.8%; $I^2 = 97\%$) of patients who entered the open label period finished the open label period (see Results S1: figure 3). The association of study duration and the logit of the event rate was $\beta = -0.002$, $p = .62$; $\tau^2 = 0.87$; $p < .0001$.

3.4.4 | Patients randomized at baseline/ finishing open label period

Fourteen studies with 3,136 participants were entered into an analysis of patients who were randomized at baseline and finished open label period. 31.1% (23.0%–40.7%; $I^2 = 97\%$) of patients who were randomized at baseline, finished the open label period (see Results S1: figure 4). The association of study duration and the logit of the event rate was $\beta = -0.02$, $p = .0002$; $\tau^2 = 0.66$; $p < .0001$.

In sum, the total loss was 68.9% (59.3%–77.0%) of all patients primarily included into the randomized controlled trial.

3.4.5 | Patients dropping out due to lack of efficacy

Eleven studies with 2012 patients were entered into an analysis of dropping out due lack of efficacy. 3.9% (2.5%–6.1%; $I^2 = 71\%$) of patients dropped out due lack of efficacy (see Results S1: figure 5). The association of study duration and the logit of the event rate was $\beta = -0.05$, $p = .0006$; $\tau^2 = 0.55$; $p < .0001$.

3.4.6 | Patients dropping out due to adverse events

Twelve studies with 3,362 participants were entered into an analysis of patients dropping out due to adverse events. 14.1% (10.0%–19.4%; $I^2 = 92\%$) of patients dropped out to due adverse events in the open label period (see Results S1:

figure 6). The association of study duration and the logit of the event rate was $\beta = -0.03$, $p < .0001$; $\tau^2 = 1.2$; $p < .0001$.

3.4.7 | Patients with serious adverse events

Twelve studies with 3,362 participants were entered into an analysis of serious adverse events. In 6.3% (3.9%–10.1%; $I^2 = 92\%$) of patients, serious adverse events were noted in the open label period (see Results S1: figure 7). The association of study duration and the logit of the event rate was $\beta = -0.04$, $p < .0001$; $\tau^2 = 1.1$; $p < .0001$.

3.4.8 | Deaths

Eight studies with 2,905 patients were entered into an analysis of death during the open label period. 0.5% (0.2%–1.4%; $I^2 = 65\%$) of patients died during the open-label period (see Results S1: figure 8). The association of study duration and the logit of the event rate was $\beta = -0.06$, $p = .007$; $\tau^2 = 1.1$; $p < .0001$.

3.4.9 | Aberrant drug behaviour

Three studies with 302 patients were entered into an analysis of aberrant drug behaviour during the open label period. In 2.7% (1.5%–4.7%; $I^2 = 0\%$) of patients signs of aberrant drug behaviour were noted during the open-label period (see Results S1: figure 9). With only three studies available we did not perform regression analysis.

3.4.10 | Withdrawal symptoms

Four studies with 1,873 patients were entered into an analysis of withdrawal symptoms during open label period. In 2.4% (1.1%–5.2%; $I^2 = 71\%$) of patients, withdrawal symptoms were noted (see Results S1: figure 10). The association of study duration and the logit of the event rate was $\beta = -0.08$, $p < .0001$; $\tau^2 = 1.4$; $p < .0001$.

3.5 | Subgroup analyses

Due to the small number of studies, only one of the predefined subgroup analyses could be performed. *P*-value of the comparison of pain intensity at the beginning and end of open-label study was 0.24 for the comparison of the two studies with buprenorphine versus the four studies with oxycodone. Drop out rate due to adverse events was 3.8% (95% CI 2.5%–5.7%; $I^2 = 0\%$) in two studies with CLBP and

was 29.9% (95% CI 24.7%–35.8%; $I^2 = 42%$ in three studies with OA pain. Event rate of serious adverse events was 4.8% (95% CI 3.4%–6.9%; $I^2 = 0%$ in two studies with CLBP and was 5.1% (95% CI 1.5%–16.1%; $I^2 = 65%$ in three studies with OA pain.

3.6 | Sensitivity analyses

Removing two studies of which we extracted outcomes from figures did not change the results of the analysis of pain intensity at the start and end of open label of patients on study medication at the end of randomized period.

3.7 | Heterogeneity

There was substantial heterogeneity of all outcomes except for the outcome aberrant drug behaviour. Subgroup analyses showed that studies with OA pain contributed to substantial heterogeneity.

3.8 | Publication bias

The Kendall tau of the Begg rank correlation test of the pooled estimate of patients entering/finishing open-label period was not significant ($\tau = 0.08$, P two-tailed = .65). The Egger intercept of the pooled estimate of patients entering/finishing open-label period was not significant (intercept = -0.02 , P two-tailed 0.99. Eighty-seven studies would have been necessary to bring the alpha value of this outcome to >0.05 .

4 | DISCUSSION

4.1 | Summary of main results

The updated review did not change the major findings of our previous review. Based on very low quality evidence, the majority of patients with CLBP, OA- and neuropathic pain who completed a RCT and entered an open label extension phase, reported a sustained reduction of pain and disability by opioids. Within the context of an open-label trial with regular assessments, opioids were rather well tolerated and safe. If assessed, event rates of withdrawal symptoms, aberrant drug behaviour and deaths were low. However, only one-third of the patients randomised at the beginning of the RCT finished the open label period. Drop out rates due to lack of efficacy or due to adverse events and event rates of serious adverse events and deaths increased with study duration.

4.2 | Overall completeness and applicability of evidence

We cannot rule out the possibility that negative study results had not been published or had been missed by our search strategy.

The applicability (external validity) of evidence is strongly limited for the following reasons:

1. All studies were conducted in research centres. No study was conducted in a primary care setting. All studies were sponsored by the manufacturer of the opioid tested.
2. Most studies excluded patients with clinically relevant internal diseases and current or previous substance abuse.
3. The studies did not assess some risks of LTOT such as sexual dysfunction (Hsieh, DiGiorgio, Fakunle, & Sadeghi-Nejad, 2018) and respiratory depression (Nagappa, Weingarten, Montandon, Sprung, & Chung, 2017).
4. The majority of the participants were middle-aged Caucasian women. Only one study was conducted in Asia, none in Africa.
5. Most studies did not clearly describe important patient characteristics, such as the duration of symptoms or use of cointerventions.
6. Other chronic pain syndromes, for example, fibromyalgia, headache and visceral pain syndromes were not included into the studies. Our findings cannot be extrapolated to any patient with CNCP:
7. Withdrawal symptoms and aberrant drug behaviour were only analysed in some studies.
8. The positive effects of opioids in long-term open label studies cannot be disentangled from uncontrolled co-therapies, from unspecific (placebo) effects because of the lack of a placebo group and from spontaneous improvement because of the lack of a no treatment group.

4.3 | Potential biases in the review process

We might have underestimated the methodological quality of some studies which might not have reported some details required for the risk of bias and treatment quality scores used. We relied on the reported data for quality assessment and did not ask authors for further details because we did not want to introduce a “response” bias. Some studies included de-novo patients. Because data for most outcomes were not reported separately, some event rates (e.g. patients finishing open label period) might be biased towards positive results of opioids.

4.4 | Agreements with other reviews and other cohort studies

We are not aware of another systematic review of long-term open-label extension studies. Surprisingly, the most recent

US systematic review on long-term efficacy and safety of opioids for chronic pain included observational studies of health insurance companies and clinical cohorts. It did not include open-label studies, open-label extension studies and long-term randomized head-to-head comparisons of opioids. The authors concluded that evidence is insufficient to determine the effectiveness of long-term opioid therapy (LTOT) for improving chronic pain and function. In addition, they stated that good- and fair-quality observational studies suggest that opioid therapy for chronic pain is associated with increased risk for overdose, opioid abuse, fractures, myocardial infarction and markers of sexual dysfunction (Chou et al., 2015).

Our review demonstrates a long-term (≥ 26 weeks) sustained reduction of pain and disability in ca 30% of patients who were initially included into a RCT. This result is in line with the one of a short-term RCT with strict inclusion criteria like failure of guideline-based pretreatment (Maier et al., 2002). The findings from open label extension studies concur also with the results of long-term case series with selected patients of clinical centres. Of 121 patients of a German pain centre, 103 (85%) still took an opioid after an average treatment time of 66 months (37–105 months; 87% more than 5 years; Maier, Schaub, Willweber-Strumpf, & Zenz, 2005).

We did not detect the risks outlined by Chou et al. (2015) within the studies analysed except a risk of abuse in two US studies which assessed this outcome. Remarkably, the highest prevalence of aberrant drug behaviour (ca 6%) has been described early for oxycodone by Portenoy et al. (2007). Despite this finding, most studies which were later conducted and published did not assess this outcome. Recently, some manufacturers of opioids were punished for misleading the public about the risks of opioids in the USA (The Guardian, 2019).

Only some new studies included in the updated review, screened for signals of abuse and physical dependence. The opioid crisis in North America started since 2000 with a sharp increase of prescriptions for extended release oxycodone (OxyContin), which was frequently prescribed because of a presumed lower likelihood of abuse, while in reality were heavily abused. Currently, the sharp increase in opioid-associated deaths since 2014 was mainly driven by illicit fentanyl and heroin and not by prescribed opioids in the USA (Volkow et al., 2019). Without any doubt, opioids—as any other centrally acting analgesic—can cause severe side effects. Overprescription of opioids by physicians to patients was a key factor of the start of the US opioid epidemic (Rose, 2018). However, the magnitude of the crisis in North America was driven by prescribing high dosages to patients with poorly defined chronic pain syndromes as a refuge from physical and psychological trauma, social disadvantage and hopelessness (Dasgupta, Beletsky, & Ciccarone, 2018). The exclusion criteria of the studies analysed, especially mental comorbidities such as a history of substance abuse and

depression, were neglected in routine clinical care in the USA (Boscarino et al., 2010).

Significant increase of prescription opioids and non-trivial increase in opioid-related morbidity-mortality has been found in France (Chenaf et al., 2019).

No aberrant drug behaviour was reported in the Japanese study analysed (Kawamata et al., 2019). No opioid crisis has been reported in Japan (Onishi et al., 2017). Opioid prescriptions, abuse of prescribed opioids and opioid-related deaths did not increase in Germany in the last 5 years (Häuser et al., 2017; Kraus et al., 2019; Seitz et al., 2019).

The harms (abuse of prescribed opioids, mortality) of LTOT in clinical practice are underestimated by long-term extension studies probably because patients with major medical diseases and previous and current substance abuse were excluded—in contrast to a substantial part of patients in the US cohort studies analysed by Chou et al. (2015). In addition, the study context may not replicate clinical practice in the intensity of monitoring, the experience of the prescriber and the willingness to respond to specific patient complaints.

Taken together, the potential benefits of LTOT for CNCP have been neglected by the CDC guidelines review by excluding open label studies as required by drug agencies from analysis. The risks of LTOT might have been overestimated by the opioid safety studies cited in the CDC guidelines (Dowell et al., 2016). Internal validity concerns were related to poor confounding control, variable misclassification, selection bias, competing risks and potential competing interventions. External validity concerns arose from the use of limited source populations and issues with handling of cancer and acute pain patients' data (Ranapurwala, Naumann, Austin, Dasgupta, & Marshall, 2019). On the other hand, open label extension studies might overestimate the benefits and underestimate the risk of opioids as outlined in Section 4.2 and 4.3.

The average dosage of opioids did not substantially increase in most studies. The average dosages reported ranged between 40 and 100 morphine equivalent/day support the recommended dosages of recent US and Canadian guidelines (Busse et al., 2017; Dowell et al., 2016). The French guidelines recommend that the mean dosage for long-term treatment should be < 100 mg equivalent/day and propose a maximum dosage of 150 mg/day with higher dosages necessitating a specialized consultation (Moisset et al., 2016). The American Society of Interventional Pain Physicians (ASIPP) guidelines stipulate that 40 morphine milligram equivalent (MME) can be considered as low dose, 41 to 90 MME as a moderate dose, and greater than 91 MME as high dose (Manchikanti et al., 2017). However, the range of dosages reported in the studies analysed demonstrate that some patients require higher dosages of opioids for a sufficient pain relief than the recommended thresholds of the guidelines mentioned above.

5 | CONCLUSIONS

A minority of carefully selected and monitored patients with chronic low back, diabetic polyneuropathy and osteoarthritis pain profit from LTOT, even for years. These findings cannot be extrapolated to other CNCP syndromes such as other neuropathic pain syndromes, fibromyalgia syndrome or chronic visceral pain syndromes. The number of patients which remain in a LTOT decreases with time. The findings of the review support the recommendations of European Pain Federation position paper on appropriate opioid use in chronic pain management that opioid therapy can be conducted in properly selected and supervised patients within a multicomponent approach (O'Brien et al., 2017).

6 | TASKS FOR FUTURE RESEARCH

Well-designed studies are urgently needed to address the long-term benefits and risks of LTOT for CNCP (Chou et al., 2015). It is unlikely, that placebo-controlled randomized trials >1 year will be conducted by pharmaceutical companies. Public funded more flexible, large pragmatic studies or well-designed controlled observational studies (claims data of health insurance companies; patient registers), with assessment of and control for potential confounders, could advance knowledge in this area (Häuser, Bock, et al., 2015; Ranapurwala et al., 2019).

Studies that include patients who are potentially at higher risk for adverse outcomes such as major medical diseases and mental disorders are needed because such patients are commonly prescribed LTOT (Chou et al., 2015).

Internationally accepted definitions of dependence and addiction of prescribed opioids are necessary (Ballantyne, Sullivan, & Kolodny, 2012). Instruments as suggested by The Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (O'Connor et al., 2013) suggested to assess misuse, abuse and related events occurring in analgesic clinical trials and postmarketing adverse event surveillance and monitoring should be used.

ACKNOWLEDGEMENT

We thank Kathrin Bernardy for her contributions to the first version of the review.

CONFLICTS OF INTEREST

PK and WH have no financial conflicts of interest to declare. WH is the head of the German guidelines group on LTOT for CNCP. BP received honoraria for educational lectures by Mundipharma, Kyowa Kirin, Grünenthal and Indivior.

CM has received honoraria for speaking and advisory board membership from Gruenthal, MSD, Köhler Chemie, Mundipharma, Pfizer and Wyeth.

AUTHOR CONTRIBUTIONS

PK and WH performed the search of literature. WH, KB and PB selected the studies. WH, KB and PB extracted data. WH entered the data into Revman. KB and PB checked the data entry. WH wrote the manuscript. All authors discussed the results and commented on the manuscript.

REFERENCES

- Ballantyne, J. C., Sullivan, M. D., & Kolodny, A. (2012). Opioid dependence vs addiction: A distinction without a difference? *Archives of Internal Medicine*, *172*, 1342–1343. <https://doi.org/10.1001/archinternmed.2012.3212>
- Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, *50*, 1088–1101. <https://doi.org/10.2307/2533446>
- Blagden, M., Hafer, J., Duerr, H., Hopp, M., & Bosse, B. (2014). Long-term evaluation of combined prolonged-release oxycodone and naloxone in patients with moderate-to-severe chronic pain: Pooled analysis of extension phases of two Phase III trials. *Neurogastroenterology and Motility*, *26*, 1792–17801. <https://doi.org/10.1111/nmo.12463>
- Boscarino, J. A., Rukstalis, M., Hoffman, S. N., Han, J. J., Erlich, P. M., Gerhard, G. S., & Stewart, W. F. (2010). Risk factors for drug dependence among out-patients on opioid therapy in a large US health-care system. *Addiction*, *105*, 1776–1782. <https://doi.org/10.1111/j.1360-0443.2010.03052.x>
- Busse, J. W., Craigie, S., Juurlink, D. N., Buckley, D. N., Wang, L., Couban, R. J., ... Guyatt, G. H. (2017). Guideline for opioid therapy and chronic noncancer pain. *CMAJ*, *189*, E659–E666.
- Buynak, R., Rappaport, S. A., Rod, K., Arsenault, P., Heisig, F., Rauschkolb, C., & Etropolski, P. (2015). Long-term safety and efficacy of tapentadol extended release following up to 2 years of treatment in patients with moderate to severe, chronic pain: Results of an open-label extension trial. *Clinical Therapeutics*, *37*, 2420–2438. <https://doi.org/10.1016/j.clinthera.2015.08.014>
- Caldwell, J. R., Rapoport, R. J., Davis, J. C., Offenberg, H. L., Marker, H. W., Roth, S. H., ... Lynch, P. M. (2002). Efficacy and safety of a once-daily morphine formulation in chronic, moderate-to-severe osteoarthritis pain: Results from a randomized, placebo-controlled, doubleblind trial and an open-label extension trial. *Journal of Pain and Symptom Management*, *23*, 278–291. [https://doi.org/10.1016/S0885-3924\(02\)00383-4](https://doi.org/10.1016/S0885-3924(02)00383-4)
- Chenaf, C., Kaboré, J. L., Delorme, J., Pereira, B., Mulliez, A., Zenut, M., ... Authier, N. (2019). Prescription opioid analgesic use in France: Trends and impact on morbidity-mortality. *European Journal of Pain*, *23*, 124–134. <https://doi.org/10.1002/ejp.1291>
- Chou, R., Turner, J. A., Devine, E. B., Hansen, R. N., Sullivan, S. D., Blazina, I., ... Deyo, R. A. (2015). The effectiveness and risks of long-term opioid therapy for chronic pain: A systematic review for a National Institutes of Health Pathways to Prevention Workshop. *Annals of Internal Medicine*, *162*, 276–286. <https://doi.org/10.7326/M14-2559>

- Cloutier, C., Taliano, J., O'Mahony, W., Csanadi, M., Cohen, G., Sutton, I., ... Michalko, K. J. (2013). Controlled-release oxycodone and naloxone in the treatment of chronic low back pain: A placebo-controlled, randomized study. *Pain Research and Management, 18*, 75–82. <https://doi.org/10.1155/2013/164609>
- Comprehensive meta-analysis software*. (2010). Version 2.0. Englewood, NJ: Biostat.
- Dasgupta, N., Beletsky, L., & Ciccarone, D. (2018). Opioid crisis: No easy fix to its social and economic determinants. *American Journal of Public Health, 108*(2), 182–186. <https://doi.org/10.2105/AJPH.2017.304187>
- Dowell, D., Haegerich, T. M., & Chou, R. (2016). CDC guideline for prescribing opioids for chronic pain—United States, 2016. *JAMA, 315*, 1624–1645. <https://doi.org/10.1001/jama.2016.1464>
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 31*, 629–634. <https://doi.org/10.1136/bmj.315.7109.629>
- Gordon, A., Callaghan, D., Spink, D., Cloutier, C., Dzungowski, P., O'Mahony, W., ... Michalko, K. J. (2010). 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, 32*, 844–860. <https://doi.org/10.1016/j.clinthera.2010.04.018>
- Hale, M. E., Zimmerman, T. R., Ma, Y., & Malamut, R. (2015). Twelve-month, open-label assessment of long-term safety and abuse potential of hydrocodone extended-release formulated with abuse-deterrence technology in patients with chronic pain. *Journal of Opioid Management, 11*, 425–434. <https://doi.org/10.5055/jom.2015.0292>
- Hale, M., Urdaneta, V., Kirby, M. T., Xiang, Q., & Rauck, R. (2017). Long-term safety and analgesic efficacy of buprenorphine buccal film in patients with moderate-to-severe chronic pain requiring around-the-clock opioids. *Journal of Pain Research, 10*, 233–240. <https://doi.org/10.2147/jpr.s120170>
- Harati, Y., Gooch, C., Swenson, M., Edelman, S. V., Greene, D., Raskin, P., ... Kamin, M. (2000). Maintenance of the long-term effectiveness of tramadol in treatment of the pain of diabetic neuropathy. *Journal of Diabetes and Its Complications, 14*, 65–70. [https://doi.org/10.1016/S1056-8727\(00\)00060-X](https://doi.org/10.1016/S1056-8727(00)00060-X)
- Häuser, W., Bock, F., Engeser, P., Hege-Scheuing, G., Hüppe, M., Lindena, G., ... Petzke, F. (2015). Recommendations of the updated LONTS guidelines. Long-term opioid therapy for chronic noncancer pain. *Schmerz, 30*, 109–130.
- Häuser, W., Bernardy, K., & Maier, C. (2015). Long-term opioid therapy in chronic noncancer pain. A systematic review and meta-analysis of efficacy, tolerability and safety in open-label extension trials with study duration of at least 26 weeks. *Schmerz, 29*, 96–108.
- Häuser, W., Schug, S., & Furlan, A. D. (2017). The opioid epidemic and national guidelines for opioid therapy for chronic noncancer pain: A perspective from different continents. *PAIN Reports, 2*, e599. <https://doi.org/10.1097/PR9.0000000000000599>
- Johnson & Johnson (2010). Tapentadol: Clinical Study Report Synopsis CR013567 (KF5503/18). Retrieved from http://filehosting.pharm.acm.com/DownloadService.aspx?client=CTR_JNJ_6051&studyxmlid=70&filename=CR013567_CSR.pdf Accessed 28 December 2013.
- Kawamata, M., Iseki, M., Kawakami, M., Yabuki, S., Sasaki, T., Ishida, M., ... Kikuchi, S. I. (2019). Efficacy and safety of controlled-release oxycodone for the management of moderate-to-severe chronic low back pain in Japan: Results of an enriched enrollment randomized withdrawal study followed by an open-label extension study. *Journal of Pain Research, 12*, 363–375. <https://doi.org/10.2147/jpr.s179110>
- Hsieh, A., DiGiorgio, L., Fakunle, M., & Sadeghi-Nejad, H. (2018). Management strategies in opioid abuse and sexual dysfunction: A review of opioid-induced androgen deficiency. *Sexual Medicine Reviews, 6*, 618–623. <https://doi.org/10.1016/j.sxmr.2018.04.003>
- Kraus, L., Seitz, N. N., Schulte, B., Cremer-Schaeffer, P., Braun, B., Verthein, U., & Pfeiffer-Gerschel, T. (2019). Estimation of the number of people with opioid addiction in Germany. *Deutsches Ärzteblatt Online, 116*, 137–143. <https://doi.org/10.3238/arztebl.2019.0137>
- Langendam, M. W., Akl, E. A., Dahm, P., Glasziou, P., Guyatt, G., & Schünemann, H. J. (2013). Assessing and presenting summaries of evidence in Cochrane Reviews. *Systematic Reviews, 2013*, 2–81.
- McIlwain, H., & Ahdieh, H. (2005). Safety, tolerability, and effectiveness of oxymorphone extended release for moderate to severe osteoarthritis pain: A one-year study. *American Journal of Therapeutics, 12*, 106–112. <https://doi.org/10.1097/01.mjt.0000139442.65914.f9>
- Maier, C., Hildebrandt, J., Klinger, R., Henrich-Eberl, C., & Lindena, G.; MONTAS Study Group. (2002). Morphine responsiveness, efficacy and tolerability in patients with chronic non-tumor associated pain - results of a double-blind placebo-controlled trial (MONTAS). *Pain, 97*, 223–233. [https://doi.org/10.1016/S0304-3959\(02\)00020-9](https://doi.org/10.1016/S0304-3959(02)00020-9)
- Maier, C., Schaub, C., Willweber-Strumpf, A., & Zenz, M. (2005). Long-term efficiency of opioid medication in patients with chronic non-cancer-associated pain. Results of a survey 5 years after onset of medical treatment. *Schmerz, 19*, 410–417.
- Manchikanti, L., Kaye, A. M., Knezevic, N. N., McAnally, H., Slavin, K., Trescot, A. M., ... Hirsch, J. A. (2017). Responsible, safe, and effective prescription of opioids for chronic non-cancer pain: American society of interventional pain physicians (ASIPP) guidelines. *Pain Physician, 20*, S3–S92.
- Moher, D., Liberati, A., Teztlaff, J., & Altman, G.; The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 51*, 1–7.
- Moisset, X., Trouvin, A. P., Tran, V. T., Authier, N., Vergne-Salle, P., Piano, V., & Martinez, V. (2016). Use of strong opioids in chronic non-cancer pain in adults. Evidence-based recommendations from the French Society for the Study and Treatment of Pain. *Presse Médicale Paris France, 45*, 447–462.
- Moore, A. R., Eccleston, C., Derry, S., Wiffen, P., Bell, R. F., & Straube, S. (2010). “Evidence” in chronic pain—establishing best practice in the reporting of systematic reviews. *Pain, 150*, 386–389.
- Nagappa, M., Weingarten, T. N., Montandon, G., Sprung, J., & Chung, F. (2017). Opioids, respiratory depression, and sleep-disordered breathing. *Best Practice & Research Clinical Anaesthesiology, 31*, 469–485. <https://doi.org/10.1016/j.bpa.2017.05.004>
- Noble, M., Treadwell, J. R., Tregear, S. J., Coates, V. H., Wiffen, P. J., Akafomo, C., & Schoelles, K. M. (2010). Long-term opioid management for chronic noncancer pain. *The Cochrane Database System Reviews, 1*, CD006605.
- O'Connor, A. B., Turk, D. C., Dworkin, R. H., Katz, N. P., Colucci, R., Haythornthwaite, J. A., ... Zaczyn, J. P. (2013). Abuse liability measures for use in analgesic clinical trials in patients with pain: IMMPACT recommendations. *Pain, 154*, 2324–2334. <https://doi.org/10.1016/j.pain.2013.06.035>
- O'Brien, T., Christrup, L. L., Drewes, A. M., Fallon, M. T., Kress, H. G., McQuay, H. J., ... Wells, J. C. (2017). European Pain Federation position paper on appropriate opioid use in chronic pain management. *European Journal of Pain, 21*, 3–19. <https://doi.org/10.1002/ejp.970>
- Onishi, E., Kobayashi, T., Dexter, E., Marino, M., Maeno, T., & Deyo, R. A. (2017). Comparison of opioid prescribing patterns in the

- United States and Japan: Primary care physicians' attitudes and perceptions. *The Journal of the American Board of Family Medicine*, 30, 248–254. <https://doi.org/10.3122/jabfm.2017.02.160299>
- Orwin, G. R. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics*, 8, 157–159. <https://doi.org/10.2307/1164923>
- Petzke, F., Welsch, P., Klose, P., Schaefer, R., Sommer, C., & Häuser, W. (2015). Opioids in chronic low back pain. A systematic review and meta-analysis of efficacy, tolerability and safety in randomized placebo-controlled studies of at least 4 weeks duration. *Schmerz*, 29, 60–72. Erratum. In: *Schmerz* (2015). 29, 308.
- Portenoy, R. K., Farrar, J. T., Backonja, M. M., Cleeland, C. S., Yang, K., Friedman, M., ... Richards, P. (2007). Long-term use of controlled-release oxycodone for noncancer pain: Results of a 3-year registry study. *The Clinical Journal of Pain*, 23, 287–299. <https://doi.org/10.1097/AJP.0b013e31802b582f>
- Ranapurwala, S. I., Naumann, R. B., Austin, A. E., Dasgupta, N., & Marshall, S. W. (2019). Methodologic limitations of prescription opioid safety research and recommendations for improving the evidence base. *Pharmacoepidemiology and Drug Safety*, 28, 4–12. <https://doi.org/10.1002/pds.4564>
- Ray, W. A., Chung, C. P., Murray, K. T., Hall, K., & Stein, C. M. (2016). Prescription of long-acting opioids and mortality in patients with chronic noncancer pain. *JAMA*, 315, 2415–2423. <https://doi.org/10.1001/jama.2016.7789>
- Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.
- Richarz, U., Waechter, S., Sabatowski, R., Szczepanski, L., & Binsfeld, H. (2013). Sustained safety and efficacy of once-daily hydromorphone extended-release (OROS® hydromorphone ER) compared with twice-daily oxycodone controlled-release over 52 weeks in patients with moderate to severe chronic noncancer pain. *Pain Practice*, 13, 30–40. <https://doi.org/10.1111/j.1533-2500.2012.00553.x>
- Rose, M. E. (2018). Are prescription opioids driving the opioid crisis? Assumptions vs facts. *Pain Medicine*, 19, 793–807. <https://doi.org/10.1093/pm/pnx048>
- Roth, S. H., Fleischmann, R. M., Burch, F. X., Dietz, F., Bockow, B., Rapoport, R. J., ... Lacouture, P. G. (2000). Around-the-clock, controlled-release oxycodone therapy for osteoarthritis-related pain: Placebocontrolled trial and long-term evaluation. *Archives of Internal Medicine*, 160, 853–860. <https://doi.org/10.1001/archinte.160.6.853>
- Sandner-Kiesling, A., Leyendecker, P., Hopp, M., Tarau, L., Lejcko, J., Meissner, W., ... Reimer, K. (2010). Long-term efficacy and safety of combined prolonged-release oxycodone and naloxone in the management of non-cancer chronic pain. *International Journal of Clinical Practice*, 64, 763–774. <https://doi.org/10.1111/j.1742-1241.2010.02360.x>
- Schaefer, R., Welsch, P., Klose, P., Sommer, C., Petzke, F., & Häuser, W. (2015). Opioids in chronic osteoarthritis pain. A systematic review and meta-analysis of efficacy, tolerability and safety in randomized placebo-controlled studies of at least 4 weeks duration. *Schmerz*, 29, 47–59. <https://doi.org/10.1007/s00482-014-1451-1>
- Seitz, N. N., Lochbühler, K., Atzendorf, J., Rauschert, C., Pfeiffer-Gerschel, T., & Kraus, L. (2019). Trends in substance use and related disorders—analysis of the Epidemiological Survey of Substance Abuse 1995 to 2018. *Deutsches Ärzteblatt International*, 116, 585–591.
- Sommer, C., Welsch, P., Klose, P., Schaefer, R., Petzke, F., & Häuser, W. (2015). Opioids in chronic neuropathic pain. A systematic review and meta-analysis of efficacy, tolerability and safety in randomized placebo-controlled studies of at least 4 weeks duration. *Schmerz*, 29, 35–46.
- Thorne, C., Beaulieu, A. D., Callaghan, D. J., O'Mahony, W. F., Bartlett, J. M., Knight, R., ... Darke, A. C. (2008). A randomized, double-blind, crossover comparison of the efficacy and safety of oral controlled-release tramadol and placebo in patients with painful osteoarthritis. *Pain Research and Management*, 13, 93–102. <https://doi.org/10.1155/2008/165421>
- Volkow, N. D., Jones, E. B., Einstein, E. B., & Wargo, E. M. (2019). Prevention and treatment of opioid misuse and addiction: A review. *JAMA Psychiatry*, 76, 208–216. <https://doi.org/10.1001/jamapsychiatry.2018.3126>
- von Korff, M., Saunders, K., Thomas Ray, G., Boudreau, D., Campbell, C., & Merrill, J., ... Weisner, C. (2008). De facto long-term opioid therapy for noncancer pain. *Clinical Journal of Pain*, 24(521–7), 1249–1255.
- Vowles, K. E., McEntee, M. L., Julnes, P. S., Frohe, T., Ney, J. P., & van der Goes, D. N. (2015). Rates of opioid misuse, abuse, and addiction in chronic pain: A systematic review and data synthesis. *Pain*, 156, 569–576. <https://doi.org/10.1097/01.j.pain.0000460357.01998.f1>
- Zeng, C., Dubreuil, M., LaRochelle, M. R., Lu, N., Wei, J., Choi, H. K., ... Zhang, Y. (2019). Association of tramadol with all-cause mortality among patients with osteoarthritis. *JAMA*, 321, 969–982. <https://doi.org/10.1001/jama.2019.1347>

WEB REFERENCES

- European Medicines Agency. (2017). Guideline on the clinical development of medicinal products intended for the treatment of pain. 01/07/2017. Retrieved from <https://www.ema.europa.eu/en/clinical-development-medicinal-products-intended-treatment-pain> Accessed April 20, 2019.
- Higgins, J. P. T., Churchill, R., Chandler, J., & Cumpston, M. S. (editors). (2017). Cochrane handbook for systematic reviews of interventions version 5.2.0 (updated February 2017), Cochrane, 2017. Retrieved from <https://training.cochrane.org/handbook>
- The Guardian. (2019). US states fight back against Purdue Pharma's bid to stop opioid lawsuits. Retrieved from <https://www.theguardian.com/us-news/2019/oct/04/purdue-pharma-oxycotin-sacklers-opioid-lawsuits>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Bialas P, Maier C, Klose P, Häuser W. Efficacy and harms of long-term opioid therapy in chronic non-cancer pain: Systematic review and meta-analysis of open-label extension trials with a study duration ≥ 26 weeks. *Eur J Pain*. 2020;24:265–278. <https://doi.org/10.1002/ejp.1496>