SYSTEMATIC REVIEW

Opioids for chronic non-cancer neuropathic pain. An updated systematic review and meta-analysis of efficacy, tolerability and safety in randomized placebo-controlled studies of at least 4 weeks duration

Claudia Sommer1 | Petra Klose2 | Patrick Welsch3 | Frank Petzke4 | Winfried Häuser3,5

1Department of Neurology, University of Würzburg, Würzburg, Germany
2Department Internal and Integrative Medicine, Faculty of Medicine, University of Duisburg-Essen, Essen, Germany
3Health Care Center for Pain Medicine and Mental Health, Saarbrücken, Germany
4Pain Medicine, Department of Anesthesiology, University Medical Center Göttingen, Göttingen, 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, tolerability and safety of opioids compared to placebo in chronic non-cancer neuropathic pain.

Databases and Data Treatment: Clinicaltrials.gov, CENTRAL, PubMed and PsycINFO were searched from October 2013 to June 2019. Randomized controlled trials comparing opioids with placebo and at least 4 weeks double-blinded duration were analysed. Primary outcomes were pain relief of 50% or greater, disability, tolerability and safety. Effects were summarized by a random effects model using risk differences (RD) or standardized mean differences (SMD). We added four new studies with 662 participants for a total of 16 included studies with 2,199 participants. Study duration ranged between 4 and 12 weeks. Studies with a parallel and crossover design: Based on low to moderate quality evidence, opioids (buprenorphine, hydromorphone, morphine, oxycodone, tramadol) provided a clinically relevant pain relief of 50% or greater and reduction of disability compared to placebo. There was no clinically relevant harm with regards to the drop out rate due to adverse and serious adverse events by opioids compared to placebo. Enriched enrolment randomized withdrawal design: Based on low to moderate quality evidence, tapentadol provided a clinically relevant pain relief of 50% or greater and reduction of disability compared to placebo in diabetic polyneuropathy. There was no clinically relevant harm with regards to the drop out rate due to adverse and serious adverse events by tapentadol compared to placebo.

Conclusions: Some opioids provided a short-term substantial pain relief in highly selected patients in some neuropathic pain syndromes.

Significance: Some opioids (buprenorphine, morphine, oxycodone, tramadol, tapentadol) provide substantial pain relief compared to placebo in postherpetic neuralgia and peripheral neuropathies of different aetiologies for 4–12 weeks. There is insufficient evidence to support or refute the suggestion that these drugs are effective in other neuropathic pain conditions. The safety of opioids with regards to abuse and deaths in the studies analysed cannot be extrapolated to routine clinical care.
A large variety of insults to the peripheral or central somatosensory nervous system may result in chronic neuropathic pain (CNP). Common examples of peripheral neuropathic pain include diabetic and postsurgical neuropathy. Central neuropathic pain includes poststroke pain, pain in multiple sclerosis and pain after spinal cord injury. Estimates of the prevalence of chronic pain with neuropathic characteristics range from 6.9% to 10% in the general population (Hecke, Austin, Khan, Smith, & Torrance, 2014).

In order to facilitate the assessment and treatment of CNP, clinical practice guidelines have been published by international pain associations such as the International Association for the Study of Pain (IASP) (Finnerup et al., 2015) and the European Federation of Neurological Societies (EFNS) (Attal et al., 2010). None of these guidelines recommended opioids (despite evidence for efficacy) as first line therapy because of concerns about safety, with tramadol as a potential second line option. In addition, some national guidelines on opioids for chronic non-cancer pain (CNCP) such as the French (Moisset & Martinez, 2016) and German guidelines (Häuser et al., 2015a) commented on the efficacy and safety of opioids for CNP. All guidelines based their recommendations on meta-analyses of randomized controlled trials (RCTs). The most up-to-date search of the literature for all opioids for CNP guidelines was conducted by the IASP and searched the literature until January 2014.

In the view of the opioid crisis in North America, concerns about prescription opioid misuse and use disorders to prescribed opioids have been raised (Manchikanti et al., 2012). The recent guideline on the clinical development of medicinal products intended for the treatment of pain of the European Medicines Agency (EMA) recommends that trials should be designed in such a way, that these phenomena can be studied (European Medicines Agency, 2017). To the best of our knowledge, withdrawal symptoms and rates/signs of abuse/addiction have not been analysed in previous reviews on opioids for CNP available.

For the revision of the German 2015 guidelines on long-term administration of opioids in chronic non-cancer pain (LONTS) (Häuser et al. 2015a), we updated our systematic review with meta-analysis of opioids for CNP (Sommer et al., 2015). The objectives of this updated review were to determine the efficacy, tolerability and safety (including risks of withdrawal symptoms and abuse) of opioids (including opioids with an additional mode of action) compared to placebo in non-cancer CNP patients of any age in randomized placebo-controlled studies of at least 4 weeks duration (titration and maintenance).
and transdermal routes. b opioids administered as abuse deterrent formulations (ADF), 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 European Medicines Agency (EMA).

Types of studies
We included fully published double-blind randomized controlled trials (RCTs) that compared opioids as defined above to placebo (pure or pseudo) for therapeutic purposes in CNP. We included both studies with a parallel and an enriched enrollment withdrawal (EERW) design. Studies with a crossover design were only included if (a), separated data from the two periods were reported or (b), data were presented which excluded statistically significant carry-over effects or (c), statistical adjustments were carried out in the case of a significant carry-over effect. Study duration had to be at least 4 weeks (titration and maintenance phase for parallel and cross-over design; double-blind withdrawal phase for EERW design). Studies had to include at least 10 patients per treatment arm.

We excluded studies with a parallel design which conducted an open-label run-in and a consecutive double-blind parallel design with responders from the open-label run-in period. We excluded studies with a maintenance or withdrawal period of less than 4 weeks duration, those with an experimental design (i.e. if the primary purpose was to study pain mechanisms and not pain relief) and studies which were only published as abstracts.

We grouped outcome measures according to the length of the double-blind (titration and maintenance): short-term (4–12 weeks), intermediate (13–26 weeks) and long-term (longer than 26 weeks).

Types of outcome measures
The selection of outcomes was based on the recommendations of the ACTINPAIN writing group of the International Association for the Study of Pain (IASP) Special Interest Group (SIG) on Systematic Reviews in Pain Relief (Moore et al., 2010), the guideline on the clinical development of medicinal products intended for the treatment of pain of the European Medicines Agency (European Medicines Agency, 2017) as well as those from the Cochrane Pain, Palliative and Supportive Care Systematic Review Group editors for reporting meta-analyses of RCTs in chronic pain (Cochrane Pain, Palliative and Supportive Care Systematic Review Group, 2012).

Primary outcomes. 1. Pain relief of 50% or greater (efficacy; dichotomous variable)
2. Patient global impression to be much or very much improved (efficacy; dichotomous variable)
3. Disability (efficacy; continuous variable)
4. Drop out rates to adverse events (tolerability; dichotomous variable)
5. Frequency of serious adverse events (safety; dichotomous variable)
6. Any death (safety; dichotomous variable)

Secondary outcomes. 1. Pain relief of 30% or greater (efficacy; dichotomous variable)
2. Pain intensity (efficacy; continuous variable)
3. Sleep problems (efficacy; continuous variable)
4. Drop out rates due to lack of efficacy (efficacy; dichotomous variable)
5. Withdrawal symptoms (safety; continuous or dichotomous variable)
6. Prescription opioid abuse/ opioid use disorder (safety; dichotomous variable)

2.1.2 | Electronic searches
We searched:
- The Cochrane Central Register of Controlled Trials (CENTRAL (from October 2013 to June 3, 2019): Most CENTRAL records are taken from bibliographic databases (mainly PubMed and Embase), but records are also derived from other published and unpublished sources, including ClinicalTrials.gov and the WHO's International Clinical Trials Registry Platform.
- PsychInfo (from October 2013 to June 3, 2019).
The search strategy for MEDLINE is outlined in Methods S1. The search was conducted by PK.

2.1.3 | Searching other resources

We searched http://www.clinicaltrials.gov (website of the US National Institutes of Health) for completed trials to April 12, 2019. The search was conducted by WH.

All authors searched bibliographies from retrieved relevant articles. Our search included all languages.

2.2 | Measures of treatment effect

The effect measures of choice were risk differences (RD) for dichotomous data and standardized mean difference (SMD) for continuous data (method inverse variance). We used a random-effect model because we assumed that the effects being estimated in the different studies are not identical, but follow some distribution. Uncertainty was expressed using 95% confidence intervals (CIs). Number needed to treat for an additional benefit (NNTBs) was calculated as the reciprocal of the absolute risk reduction (ARR). For unwanted effects, the NNTB becomes the number needed to treat for an additional harm (NNTH) and is calculated in the same manner. For dichotomous data, we calculated risk differences (RDs). The threshold for ‘clinically relevant benefit’ or ‘clinically relevant harm’ was set for categorical variables by an absolute risk reduction or increase ≥ 10% corresponding a NNTB or NNTH of < = 10 (Moore, Barden, Derry, & McQuay, 2008).

Cohen’s categories were used to evaluate the magnitude of the effect size of continuous data, calculated by SMD, with values for Hedges’ $g$ as follows:

0.2–0.5 equating to a small effect size, 0.5–0.8 equating to a medium effect size and more than 0.8 equating to a large effect size (Cohen, 1988). We considered values of $g$ less than 0.2 to equate to a ‘not substantial’ effect size (Sommer et al., 2015). The threshold ‘clinically relevant benefit’ was set for continuous variables by an effect size of more than 0.2 (Fayers & Hays, 2014).

2.3 | Data collection and analysis

2.3.1 | Selection of studies

Two review authors (WH, PW) independently scrutinized all the titles and abstracts and selected studies based on inclusion and exclusion criteria.

2.3.2 | Data extraction and management

Using standardized forms, three pairs of authors (CS, WH; FP, WH; PW, WH) independently extracted data on inclusion and exclusion criteria of studies, participant characteristics, intervention group, clinical setting, interventions, country of study and study sponsorship. If data were not available in a format that was appropriate for data extraction, we did not contact the authors of the trial for further clarification. Any disagreements were resolved through discussion.

2.3.3 | Assessment of risk of bias in included studies

Two review authors (PW, WH) independently assessed the risk of bias of each included trial. Disagreements were resolved by discussion and consensus, otherwise a third review author (FP) acted as arbiter. We assessed the following risks of bias for each study in accordance with methods recommended by The Cochrane Collaboration (Higgins et al., 2017): selection biases (Random sequence generation; allocation concealment; group similarity at baseline), performance bias (Blinding of participants and personnel); detection bias (blinding of outcome assessor), attrition bias (Incomplete outcome data), reporting bias (selective outcome reporting), performance bias and other bias (sample size). For details see Methods S2.

We defined a high-quality study (low risk of bias) as one that fulfilled six to eight, a moderate-quality study (moderate risk of bias) as one that fulfilled three to five and a low-quality study (high risk of bias) as one that fulfilled zero to two of the eight validity criteria.

See Methods S3 for:

2.4.5 Unit of analysis issues.
2.4.6 Dealing with missing data.
2.4.7 Assessment of heterogeneity.
2.4.8 Grading of evidence.

2.4 | Subgroup analysis

Subgroups were planned a priori to assess the variations in effect size (heterogeneity) for all types of opioids pooled together compared to placebo groups pooled together, for different types of chronic neuropathic pain, different types of opioids (pure opioids vs. opioids with additional modes of action, i.e. tramadol, tapentadol), treatment duration (short-term, intermediate-term and long-term studies) for the primary outcomes. At least two studies had to be available for subgroup analysis.
2.5 | Sensitivity analysis

We planned to perform sensitivity analysis for all types of opioids pooled together compared to placebo groups pooled together for pain relief of 50% or more in studies in which we extracted means and/or SDs from figures or calculated SDs from p-values or used imputation methods to calculate these outcomes.

2.6 | Publication bias

We assessed publication bias using a method designed to detect the amount of unpublished data with a null effect required to make one major outcome of efficacy irrelevant (usually taken to mean an NNTB of 10 or higher) (Moore et al., 2008).

2.7 | Software

RevMan Analysis (RevMan 5.3.1) software of the Cochrane Collaboration were used for statistical analyses (Review Manager (RevMan) [Computer program] 2014).

3 | RESULTS

3.1 | Search

The total number of included studies in the 2015 review was 12. The updated searches (last performed April 2019) produced 12,334 records after duplicates were removed. Three studies were excluded. One study with morphine for phantom limb pain with a cross-over design did not report that there was no carry-over effect (Huse, Larbig, Flor, & Birbaumer, 2001). One study (Arai, Kashimoto, Ukyo, Tominaga, & Imanaka, 2015) with fentanyl for neuropathic pain was excluded for two reasons: The study included patients with CRPS, but did not report the types of CRPS. The study included patients with postoperative pain without reporting that this pain was exclusively neuropathic. One study with tapentadol (as active comparator) in chronic low back pain was excluded, because the authors did not present the outcomes for the subgroup of patients with clinically diagnosed lumbar radiculopathy, but for patients with a high probability of neuropathic pain as assessed in a screening questionnaire (Christoph, Eerdekens, Kok, Volkers, & Freynhagen, 2017).

We included 12 studies from our previous review (Boureau, Legallicier, & Kabir-Ahmadi, 2003; Gilron et al., 2005; Gimbel, Richards, & Portenoy, 2003; Harati et al., 1998; Khoromi, Cui, Nackers, & Max, 2007; Norbrink & Lundeberg, 2009; Raja et al., 2002; Schwartz et al., 2011; Sindrup et al., 1999; Watson & Babul, 1998; Watson, Moulin, Watt-Watson, Gordon, & Eisenhoffer, 2003; Wu et al., 2008). We included one study (Sindrup et al., 2012) which we have missed in the previous review. We included three new studies (Nalamachu, Hale, & Khan, 2014; Simpson & Wlodarczyk, 2016; Vinik et al., 2014) and thus a total of 16 studies with 2,199 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 Table S1.

3.2.1 | Settings

Ten studies were conducted in the USA, five studies in Europe and one in Australia.

3.2.2 | Types of opioids

Four studies each tested morphine extended release and tramadol orally. Two studies each tested oxycodone extended release, tramadol and tapentadol extended release orally. One study each tested buprenorphine transdermal and hydromorphone extended release. One study used tramadol (and placebo) as comparators to a drug under development. All studies used flexible dosages. The average dosages reported for hydromorphone was 37.0 mg/d, for morphine ranged from 45 mg/d to 112 mg/d, for oxycodone ranged from 37 mg/d to 45 mg/d and for tramadol ranged from 210 mg/d to 275 mg/d. The maximum dosages allowed were 64 mg/d for hydromorphone, 180 mg/d for morphine, 240 mg/d for oxycodone and 400 mg/d for tramadol. The studies with buprenorphine (5 to 40 ug/h) and tapentadol (200 to 500 mg/d) did not report on the average dosages.

3.2.3 | Study design

Eight studies used a cross-over design, five studies a parallel design and three studies an EERW design. Double-blind phase of all studies ranged between 4 and 12 weeks.

3.2.4 | Types of CNP

Six studies included patients with painful diabetic polyneuropathy, three studies with postherpetic neuralgia and two studies with painful peripheral neuropathy of various
aetiologies (diabetic, alcoholic, drugs, unknown). Two studies each included patients with pain due lumbar root lesion and one study each with spinal cord injury and phantom limp pain.

3.2.5 | Participants

The percentage of women in the studies ranged between 22% and 100%. All studies included only adults. If reported, the mean age of the participants ranged between 40 and 68 years. If reported, the percentage of Caucasians ranged between 79% and 97%. Four studies included less than 50 participants, nine included 50–150 and three included more than 150 participants per treatment arm at study entry.

3.2.6 | Exclusion of clinically relevant somatic disease or mental disorder

Twelve studies excluded patients with relevant somatic diseases and 14 studies excluded patients with mental disorders.

FIGURE 1 PRISMA Flow Diagram
## Table 1
Overview of the randomized controlled trials in chronic neuropathic pain included into the systematic review (grouped by type of opioid in alphabetical order)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
<th>Countries of study centers</th>
<th>Study design</th>
<th>Population type</th>
<th>Exclusion of patients with other internal diseases</th>
<th>Exclusion of patients with current or previous substance abuse and addiction</th>
<th>Number of patients randomized</th>
<th>Interventions and control group</th>
<th>Duration of trial (titration and maintenance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buprenorphine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simp &amp; Wlodarczyk, 2016</td>
<td>Australia</td>
<td>Parallel</td>
<td></td>
<td>Diabetic neuropathy pain</td>
<td>No</td>
<td>No</td>
<td>186</td>
<td>Buprenorphine flexible 5,10,20,30 and 40 ug/h transdermal Placebo transdermal</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1-week screening phase, 12-week maintenance phase 3-week blinded downward titration completion phase</td>
</tr>
<tr>
<td>Hydromorphone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nalamachu et al., 2014</td>
<td>USA</td>
<td>Enriched enrollment randomized withdrawal</td>
<td></td>
<td>Neuropathic low back pain</td>
<td>Yes</td>
<td>Yes</td>
<td>94</td>
<td>Hydromorphone extended release up to 64 mg/d flexible oral Placebo oral</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 to 4 weeks open label 12 weeks double-blind withdrawal</td>
</tr>
<tr>
<td>Morphine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gilron et al., 2005</td>
<td>Canada</td>
<td>Cross over</td>
<td></td>
<td>Postherpetic neuralgia and diabetic neuropathy pain</td>
<td>Yes</td>
<td>Yes</td>
<td>57</td>
<td>Morphine sustained release up to 120 mg/d flexible oral Active placebo (lorazepam) flexible oral</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No reports on wash-out periods 4 weeks each</td>
</tr>
<tr>
<td>Khoromi et al., 2007</td>
<td>USA</td>
<td>Cross over</td>
<td></td>
<td>Lumbar root pain</td>
<td>Yes</td>
<td>Yes</td>
<td>55</td>
<td>Morphine sustained release flexible 15-90 mg/d oral Benzotropine flexible (0.25-1 mg/d) (active placebo) oral</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Duration screening and wash-out not reported 3 weeks titration 2 weeks maintenance 2 weeks tapering</td>
</tr>
<tr>
<td>Raja et al., 2002</td>
<td>USA</td>
<td>Cross over</td>
<td></td>
<td>Postherpetic neuralgia pain</td>
<td>Yes</td>
<td>Yes</td>
<td>50</td>
<td>Morphine controlled release flexible 15-240 mg/d oral Placebo oral flexible</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 week washout Approximately 8 weeks titration, maintenance and dose tapering for each period 1 week drug free wash-out</td>
</tr>
<tr>
<td>Wu et al., 2008</td>
<td>USA</td>
<td>Cross over</td>
<td></td>
<td>Postamputation (stump and phantom pain)</td>
<td>Yes</td>
<td>Yes</td>
<td>60</td>
<td>Morphine sustained release 15-180 mg/d flexible oral Oral placebo</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Wash-out time not reported 4 weeks titration 2 weeks maintenance 2 weeks dose tapering for each period 1 week drug free wash-out</td>
</tr>
<tr>
<td>Oxycodone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gimbel et al., 2003</td>
<td>USA</td>
<td>Parallel</td>
<td></td>
<td>Diabetic polyneuropathy pain</td>
<td>Yes</td>
<td>Yes</td>
<td>159</td>
<td>Oxycodone extended release oral flexible up to 20 mg/d oral Placebo oral</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Screening 7 days 6 weeks titration and maintenance 1-week tapering optional</td>
</tr>
<tr>
<td>Watson &amp; Babul, 1998</td>
<td>USA</td>
<td>Cross over</td>
<td></td>
<td>Postherpetic neuralgia</td>
<td>No</td>
<td>Yes</td>
<td>50 (evaluated 38)</td>
<td>Oxycodone controlled release oral flexible 20-160 mg/d Placebo oral</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Duration wash out and screening not reported 4 weeks titration and maintenance each No wash out between periods</td>
</tr>
</tbody>
</table>

(Continues)
<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
<th>Countries of study centers</th>
<th>Study design</th>
<th>Population type</th>
<th>Number of patients randomized</th>
<th>Interventions and control group</th>
<th>Duration of trial (titration and maintenance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watson et al., 2003</td>
<td>USA</td>
<td></td>
<td>Cross-over</td>
<td>Diabetic polyneuropathy pain</td>
<td>Yes</td>
<td>Oxycodone controlled release oral flexible 20-80 mg/d oral</td>
<td>2-7 days wash-out 4 weeks titration and maintenance each</td>
</tr>
<tr>
<td>Schwartz et al., 2011</td>
<td>Canada, USA</td>
<td></td>
<td>Enriched rollment randomized withdrawal</td>
<td>Painful diabetic polyneuropathy</td>
<td>Yes</td>
<td>Tapentadol extended release 200-500 mg/d flexible oral</td>
<td>Duration screening not reported 3 weeks open label titration phase 12 week double blind withdrawal</td>
</tr>
<tr>
<td>Vinik et al., 2014</td>
<td>Canada, USA</td>
<td></td>
<td>Enriched rollment randomized withdrawal</td>
<td>Painful diabetic polyneuropathy</td>
<td>Yes</td>
<td>Tapentadol extended release 200-500 mg/d flexible oral</td>
<td>13-day screening period 5-day washout period 3-day pretitration pain intensity evaluation period 3-week open label titration period 12 weeks double-blind withdrawal</td>
</tr>
<tr>
<td>Harati et al., 1998</td>
<td>USA</td>
<td></td>
<td>Parallel</td>
<td>Diabetic polyneuropathy</td>
<td>Yes</td>
<td>Tramadol flexible 100-400 mg/d oral</td>
<td>Duration of wash-out not reported 6 weeks titration and maintenance 6 months open label</td>
</tr>
<tr>
<td>Norbrink &amp; Lundeberg, 2009</td>
<td>Sweden</td>
<td></td>
<td>Parallel</td>
<td>Neuropathic pain after spinal cord injury</td>
<td>No</td>
<td>Tramadol flexible 100-400 mg/d oral</td>
<td>Duration of wash-out not reported 6 weeks titration and maintenance 6 months open label</td>
</tr>
<tr>
<td>Sindrup et al., 1999</td>
<td>Denmark</td>
<td></td>
<td>Cross-over</td>
<td>Polyneuropathy of different etiologies</td>
<td>Yes</td>
<td>Tramadol flexible 100-400 mg/d oral</td>
<td>Up to 1 week wash-out 4 weeks each period 1 week wash-out</td>
</tr>
<tr>
<td>Sindrup et al., 2012</td>
<td>Denmark</td>
<td></td>
<td>Cross-over</td>
<td>Polyneuropathy of different etiologies</td>
<td>Yes</td>
<td>Tramadol flexible 100-400 mg/d oral</td>
<td>Up to 1 week wash-out 4 weeks each period 1 week wash-out</td>
</tr>
</tbody>
</table>
Thirteen studies excluded patients with a history of or current substance abuse or dependence.

### 3.2.7 Funding and conflicts of interest

Ten studies reported sponsoring by pharmaceutical companies and five studies received public funding. One study did not report the details of funding.

Ten authors did not report their conflicts. Five authors declared their conflicts of interest. One author group reported that they have no conflict of interest.

### 3.3 Risk of bias in included studies

According to the predefined categories, eight studies were moderate-quality studies (unclear risk of bias overall) and eight studies were low-quality studies (high risk of bias overall) (see Figure 2 for risk of bias graph and table S2 for details).

### 3.4 Effects of intervention

#### 3.4.1 Opioids versus placebo in studies with a parallel or cross-over design at the end of treatment

**Primary outcomes**

Pain relief of 50% or greater: The outcome was calculated by an imputation method for eight studies. Eleven studies with 1,161 participants were entered into analysis. Two hundred thirty-six out of 590 (40.0%) with opioids and 123 out of 571 (21.5%) with placebo reported pain relief of 50% or greater.

RD was 0.19 [95% CI 0.13 to 0.25] ($I^2 = 28, p < .0001$). NNTB was 5 (95% CI 4 to 8). According to the predefined categories, there was a clinically relevant benefit by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Patient global impression to be much or very much improved: Only one study reported this outcome. Four out of 23 patients with tramadol, and 0 out of 12 patients with placebo reported to be much or very much improved ($p = .07$). The quality of evidence was low (downgraded by two levels due to indirectness and imprecision).

Disability: Eight studies with 861 participants were entered into analysis. SMD was −0.24 (95% CO −0.38 to −0.11) ($I^2 = 0%; p = .0004$). According to the predefined categories, the effect size was small and there was a clinically relevant benefit by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Withdrawal due to adverse events: Twelve studies with 1,339 patients were entered into analysis. One hundred and two out of 685 (14.9%) participants with opioids and 28 out of 654 (4.3%) with placebo dropped out due to adverse events. RD was 0.09 (95% CI 0.06 to 0.12) ($I^2 = 27%; p < .0001$). NNTH was 11 (95% CI 8 to 16). According to the predefined categories, there was no clinically relevant harm by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Serious adverse events: Five studies with 654 participants were entered into analysis. In 22 out of 329 (6.4%) patients with opioids and 21 out of 325 (6.5%) patients with placebo a serious adverse event was noted. RD was 0.01 (95% CI 0.03 to 0.04) ($I^2 = 22%; p = .71$). The quality of evidence was low (downgraded by two levels due to indirectness and imprecision). The study (RD 0.01 [95% CI −0.01 to 0.03]) ($I^2 = 0%; p = .31$). The quality of evidence was low (downgraded by two levels due to indirectness and imprecision).

Deaths: Only two studies with 345 participants reported explicitly this outcome. One out of 171 patients with opioids and none out of 174 patients with placebo died during the study (RD 0.01 [95% CI 0.00 to 0.03]) ($I^2 = 0%; p = .31$). The quality of evidence was low (downgraded by two levels due to indirectness and imprecision).

**Secondary outcomes**

Pain relief of 30% or greater: The outcome was calculated by an imputation method for six studies. Twelve studies with 1,226 participants were entered into analysis. Three hundred eighty-six out of 624 (61.9%) with opioids and 213 out of 602 (35.4%) with placebo reported pain relief of 30% or greater. RD was 0.28 (95% CI 0.20 to 0.36) ($I^2 = 58, p < .0001$). NNTH was 4 (95% CI 3 to 5). According to the predefined categories, there was a clinically relevant benefit by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Mean pain intensity: 13 studies with 1,325 participants were entered into analysis. SMD was −0.57 (95% CO −0.75 to −0.39) ($I^2 = 61%; p < .0001$). According to the predefined categories, the effect size was moderate and there was a clinically relevant benefit by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Sleep problems: Six studies with 666 participants were entered into analysis. SMD was −0.34 (95% CO −0.51 to −0.17) ($I^2 = 14%; p < .0001$). According to the predefined categories, the effect size was small and there was a clinically relevant benefit by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Withdrawal due to lack of efficacy: Nine studies with 1,348 participants were entered into analysis. Twenty-two out of 678 (3.2%) with opioids and 82 out of 670 (12.2%) with placebo dropped out due to lack of efficacy. RD was −0.06 (95% CI −0.10 to −0.02) ($I^2 = 70, p = .008$). NNTH was 16 (95% CI 10 to 50). According to the predefined categories,
there was no clinically relevant benefit by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Withdrawal symptoms: Only one study reported on withdrawal symptoms: In one of 36 patients with oxycodone withdrawal symptoms were noted. The quality of evidence was low (downgraded by two levels due to indirectness and imprecision [low event rate]).

Abuse and opioid use disorder: None of the studies assessed this outcome.

3.4.2 | Studies with an EERW design

**Primary outcomes**

Pain relief of 50% or greater: Two studies with 706 participants were entered into analysis. One hundred forty one out of 362 (39.0%) with tapentadol and 97 out of 344 (28.4%) with placebo reported pain relief of 50% or greater. RD was 0.11 [95% CI 0.04 to 0.18] ($I^2 = 0$, $p < 0.002$). NNTB was 9 (95% CI 6 to 25). According to the predefined categories, there was a clinically relevant benefit by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Patient global impression to be much or very much improved: Two studies with 706 participants were entered into analysis. Two hundred and fifteen out of 330 (65.2%) with tapentadol and 131 out of 316 (41.4%) with placebo reported to be much or very much improved. RD was 0.24 [95% CI 0.16 to 0.31] ($I^2 = 0$, $p < .0021$. NNTB was 4 (95% CI 3 to 6). According to the predefined categories, there was a clinically relevant benefit by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Disability: Three studies with 801 participants were entered into analysis. SMD was $-0.34$ (95% CI $-0.55$ to $-0.24$) ($I^2 = 0\%$; $p = .0001$). According to the predefined categories, the effect size was small and there was a clinically relevant benefit by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Withdrawal due to adverse events: Three studies with 801 participants were entered into analysis. Fifty-five out of 405 (13.6%) participants with opioids and 20 out of 396 (5.1%) with placebo dropped out due to adverse events, RD was 0.06 (95% CI 0.02 to 0.10) ($I^2 = 0\%$; $p = 0.0002$). NNTTH was 16 (95% CI 10 to 50). According to the predefined categories, there was no clinically relevant harm by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Serious adverse events: Three studies with 801 participants were entered into analysis. In twenty out of 405 (4.9%) patients with opioids and 17 out of 396 (4.3%) patients with placebo a serious adverse event was noted. RD was 0.01 (95% CI $-0.04$ to 0.05) ($I^2 = 49\%$; $p = .81$). The quality of evidence was moderate (downgraded by one level due to indirectness).

Death: Two studies with 414 participants were entered into analysis. One out of 211 participants (0.5%) with opioids...
and none out of 203 participants with placebo died during the study. RD was 0.01 (95% CI −0.01 to 0.02) (I² = 0%; p = .51). The quality of evidence was low (downgraded by one level due to indirectness and imprecision [low event rate]).

Secondary outcomes
Pain relief of 30% or greater: Two studies with 706 participants were entered into analysis. One hundred and ninety-seven out of 362 (54.4%) with opioids and 150 out of 344 (43.6%) with placebo reported pain relief of 30% or greater. RD was 0.11 [95% CI 0.09 to 0.18] (I² = 0, p = .004). According to the predefined categories, there was a clinically relevant benefit by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Mean pain intensity: Three studies with 799 participants were entered into analysis. SMD was −0.55 (95% CI −0.71 to −0.39) (I² = 19%; p < .0001). According to the predefined categories, the effect size was moderate and there was a clinically relevant benefit by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Sleep problems: This outcome was not assessed by the three studies.

Withdrawal due to lack of efficacy: Three studies with 799 participants were entered into analysis. Nineteen out of 405 (4.7%) with opioids and 49 out of 394 (12.4%) with placebo dropped out due to lack of efficacy. RD was −0.07 [95% CI −0.12 to −0.02] (I² = 70, p = .008). NNTB was 9 (95% CI 5 to 611). According to the predefined categories, there was a clinically relevant benefit by opioids. The quality of evidence was moderate (downgraded by one level due to indirectness).

Withdrawal symptoms: Three studies with 620 participants were entered into analysis. Nineteen out of 313 (3.5%) with opioids and 12 out of 313 (3.8%) with placebo reported withdrawal symptoms after discontinuation of medication. RD was −0.01 [95% CI −0.04 to −0.06] (I² = 56, p = .61). The quality of evidence was moderate (downgraded by one level due to indirectness).

Abuse and opioid use disorder: None of the studies assessed this outcome.

3.5 | Subgroup analyses

3.5.1 | Different types of opioids
In studies with a parallel and cross-over design, the test for subgroup differences (morphine, oxycodone, tramadol) yielded these results: Pain relief of 50% or greater: Chi² = 1.48, p = .48; Disability: Chi² = 0.75, p = .75; drop out due to adverse events Chi² = 1.61, p = .45. Chi² was 3.92 and p = .05 for the test of subgroup differences between oxycodone and tramadol for serious adverse events.

3.5.2 | Pure opioids versus opioids with an additional mode of action
In studies with a parallel and cross-over design, RD of pain relief of 50% or greater was 0.18 (95% CI 0.08 to 0.28) (I² = 51, p = .004) for morphine and oxycodone and 0.22 (95% CI 0.14 to 0.31) (I² = 0%, p < .0001) for tramadol. SMD for disability was −0.26 (95% CI 0.45 to −0.07) (I² = 0%; p = .007) for oxycodone and morphine and −0.36 (95% CI −0.67 to −0.05) (I² = 0; p = .002) for tramadol. RD for drop out rates due to adverse events was 0.07 (95% CI 0.03 to 0.11) (I² = 0%; p = .008) for morphine and oxycodone and 0.09 (95% CI 0.04 to 0.14) (I² = 11%; p = .0002) for tramadol. The differences between the subgroups were not clinically relevant.

The other predefined subgroup analyses were not possible due to the lack of data (at least two studies available for comparison).

3.6 | Sensitivity analyses
Removing the studies with imputed rates of pain relief of 50% or greater in studies with a parallel and cross-over resulted in a RD 0.18 (95% CI 0.09 to 0.27).

3.7 | Publication bias
Studies with 774 participants with a null effect on pain relief of 50% or greater or greater would have been required to make the result clinically irrelevant (NNTB of 10 or higher).

4 | DISCUSSION

The updated review did not change the major findings of our previous review. Based on low to moderate evidence, opioids (buprenorphine, hydromorphone, morphine, oxycodone, tramadol) provided in studies with a parallel and cross over with 4–12 weeks duration, in which patients were excluded if they have mental health disorders including past history of substance use disorders, a clinically relevant pain relief of 50% or greater and reduction of disability compared to placebo. There was no clinically relevant harm with regards to the drop out rate due to adverse and serious adverse events by opioids compared to placebo. Based on low to moderate evidence, in enriched enrolment randomized withdrawal design tapentadol provided a clinically relevant pain relief of 50% or greater and reduction of disability compared to placebo in diabetic polyneuropathy.
There was no clinically relevant harm with regards to the drop out rate due to adverse and serious adverse events by tapentadol compared to placebo.

### 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 limited for the following reasons:

1. Most studies were sponsored by the manufacturer of the drug tested. One sponsor of some trials will establish a nearly $200 million endowment at the Oklahoma State University's Center for Wellness and Recovery, which will go towards treating the ongoing addiction epidemic in USA. (Mike Hunter Oklahoma Attorney General., 2019)

2. Most studies were conducted in research centres. No study was conducted in a primary care setting.

3. Most studies excluded patients with clinically relevant somatic diseases and current or previous substance abuse. Somatic and mental comorbidities in patients with chronic pain are prevalent in the general population (Häuser, Schmutzer, Hilbert, Brähler, & Henningsen, 2015b).

4. The majority of the participants were middle-aged Caucasian women. Only one study was conducted in Asia, none in Africa.

5. No study included children and adolescents or presented a subgroup analysis of seniors.

6. The studies analysed to not allow to make conclusions on the long-term (more than 6 months) efficacy and safety of opioids for chronic neuropathic pain. EMA recommends open label extension studies to assess long-term efficacy and safety (European Medicines Agency, 2017). One of the studies analysed which tested tramadol for painful diabetic polyneuropathy included a 6 months open label extension phase with 117 patients which demonstrated a sustained pain relief, a low drop out rate due to lack of efficacy (3.4%) and moderate drop out rate due to adverse events (11.1%). In 10 patients serious adverse events were noted which were judged not to be treatment-related by the investigators (Harati et al., 2000). In a randomized controlled trial with transdermal buprenorphine and fentanyl for HIV polyneuropathy in 40 patients, both drugs induced a significant pain relief up to 150 days of treatment. Drop out rate due to adverse events was 0% and for lack of efficacy was 7.5%. Serious adverse events were not reported (Canneti et al., 2013).

7. Abuse of prescribed opioids was not assessed.

### 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. We used imputation methods if the rates of a moderate and substantial pain relief were not reported.

### 4.4 Agreements with other reviews

Our review suggests a clinically relevant benefit and overall safety of opioids for CNP. However, our data do not support on overall efficacy of all opioids for any chronic neuropathic syndrome for these reasons:

a. Some neuropathic pain syndromes such as central pain in multiple sclerosis or after stroke were not included in any study analysed. In one study each with chronic lumbar root and postamputation neuropathic pain opioids were not superior to placebo for pain relief. The most robust evidence in terms of number of studies and patients, different study designs, study duration and results for efficacy and safety is for oxycodone, tramadol and tapentadol for painful diabetic polyneuropathy.

b. The majority of studies used a cross-over design. Cross-over trials in which the results are not analysed separately by sequence of group are of limited scientific value (Wellek & Blettner, 2012). Only a minority of studies analysed solely the data obtained during the first trial period.

c. Some studies had a sample size for analysis < 50 participants leading to a small sample bias (Moore, Derry, & Wiifen, 2013). In contrast, the EERW design of the studies with tapentadol for painful diabetic neuropathic polyneuropathy has been suggested to be useful in the early phases of drug testing in humans (Moore et al., 2013). In addition, both studies included > 150 participants per treatment arm.

d. EMA requires that a sustained therapeutic effect in chronic pain should in general be demonstrated in pivotal efficacy trials with a treatment period of at least 12 weeks, excluding titration period (European Medicines Agency, 2017). Only two EERW studies met this criterion. The remaining studies had a double-blind duration of 4–6 weeks.

e. To justify a general indication for the treatment of neuropathic pain, efficacy needs to be demonstrated independently in both central and peripheral neuropathic pain (European Medicines Agency, 2017). None of the opioids analysed met this criterion.
Our concerns with regards to methods concur with the ones of recent Cochrane reviews of single opioids for chronic neuropathic pain. Stannard et al. (2016) analysed one study with hydromorphone in 94 patients with neuropathic LBP. They found insufficient evidence to support or refute the efficacy of hydromorphone in any neuropathic pain condition. Derry et al. (2016) analysed one study with fentanyl in 163 patients with mixed neuropathic conditions which was excluded from our analysis. They concluded that there is insufficient evidence to support or refute the suggestion that fentanyl works in any neuropathic pain condition. Gaskell, Moore, Derry, and Stannard (2016) analysed the same five RCTs with oxycodone as we did. They found only very low quality evidence that oxycodone was of value in the treatment of painful diabetic neuropathy or postherpetic neuralgia. There was no evidence for other neuropathic pain conditions. Cooper et al. (2017) analysed four small studies with morphine of which two were excluded by our review. They reasoned that the studies did not provide a reliable indication of a likely effect due to their very low quality. We included five of six studies with tramadol reviewed by Duehmke et al. (2017). The authors concluded that the information about the use of tramadol in neuropathic pain is coming from small, largely inadequate studies with potential risk of bias. That bias would normally increase the apparent benefits of tramadol.

No study assessed aberrant drug behaviour although the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA) emphasize collecting data on the occurrence of abuse, misuse and diversion, in phase 3 clinical trials of centrally acting drugs and analgesic clinical studies (Smith et al., 2017). Therefore we do not know if opioid prescription abuse occurred in the studies. The short duration of the RCT, the close surveillance of the patients and the exclusion of patients with a history of or current substance abuse might have reduced the risk of aberrant drug behaviour in the studies analysed.

5 | CONCLUSIONS

5.1 | Implications for clinical practice

Our systematic review gives no guidance for clinicians on first, second or third line drug therapies for CNP. There are only a few head-to-head comparisons of different drug classes for CNP available. A systematic review conducted in 2015 analysed morphine compared to antidepressants (two studies), an anticonvulsant (one study) and the antiarrhythmic mexilitene (one study) in different neuropathic pain syndromes and found no statistically significant differences between opioids and other drugs with regards to the reduction of pain and disability and to tolerability (Welsch, Sommer, Schiltenwolf, & Häuser, 2015). Based on GRADE ratings on efficacy, tolerability and safety as well as on values, preferences and costs, the Special Interest Group on Neuropathic Pain (NeuPSIG) of the IASP gave a strong recommendation for use as first-line treatment in neuropathic pain for tricyclic antidepressants, serotonin-noradrenaline reuptake inhibitors, pregabalin and gabapentin; a weak recommendation for use as second line for lidocaine patches, capsaicin high-concentration patches and tramadol; and a weak recommendation for use as third line for strong opioids and botulinum toxin A (Finnerup et al., 2015). In addition, the selection of drugs should be based on the comorbidities of the patient which might determine some contraindications for single drug classes and patients’ preferences, for example the subjective importance of some frequent side effects such as weight gain and sexual dysfunction (Häuser et al. 2015a).

Recent evidence-based guidelines recommended to restrict the dosage for long-term opioid therapy to 90 mg morphine equivalent (MEQ)/d (Busse et al., 2017; Dowell, Haegerich, & Chou, 2016), 120 mg MEQ/d (Häuser et al. 2015a) and 150 mg MEQ/d (Moisset & Martinez, 2016). The average dosages reported for morphine, oxycodone and tramadol in the studies analysed in this review support the lower recommended dosages. However, the range of dosages reported demonstrates that some patients require higher dosages of opioids for a sufficient pain relief than the recommended thresholds of the guidelines mentioned above.

Risk factors for prescription opioid abuse such as a history of substance abuse should be investigated before prescription of opioids and signals of misuse should be assessed at each renewal (Häuser et al., 2015; Moisset & Martinez, 2016).

5.2 | Major tasks for future research

Studies should follow EMA recommendations (European Medicines Agency, 2017) for the design of chronic pain trials:

a. If there is an established treatment option a three-arm study (study drug—comparator—placebo) should be provided in order to allow the assessment of comparative efficacy and safety of a new product.

b. Double-blind phase should be at least 12 weeks.

c. Opioid withdrawal symptoms (physical dependence) after the end of study and aberrant drug behaviour (abuse, opioid use disorder) during the study should be assessed. The Abuse Liability Evaluation for Research, Treatment, and Training (ALERTT) working group of the Analgesic, Anesthetic, and Addiction Clinical Trial Translations, Innovations, Opportunities, and Networks (ACTTION) has recommended to use the Self-Reported Misuse, Abuse, and Diversion [SRMAD] instrument and the Misuse, Abuse, and Diversion Drug Event Reporting System [MADDERS] to assess inappropriate medication use events (i.e. misuse, abuse and related events occurring in clinical trials (Smith et al., 2017).
d. Studies with EERW designs should use appropriate outcome measures such as loss (or maintenance) of therapeutic response.

e. The methods to assess subjective adverse events (spontaneous reports; open questions; standardized questionnaires) should be defined by drug agencies and fully reported in the publications.

f. Patients with peripheral neuropathic pain can be grouped into three sensory phenotypes based on quantitative sensory testing profiles (Vollert et al., 2017). Inclusion criteria or analyses of based on phenotypes might provide more reliable results.

**CONFLICTS OF INTEREST**

FP, PK and WH have no financial conflicts of interest to declare. FP and WH are members of the German guidelines group on long-term opioid therapy for chronic non-cancer pain. CS has received honoraria for consulting or lectures from Air Liquide, Astellas, Grünenthal and Pfizer. She has taken part in clinical trials on neuropathic pain sponsored by Biogen, Novartis andVertex and on complex regional pain syndrome sponsored by Grünenthal.

**AUTHOR CONTRIBUTIONS**

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

**REFERENCES**


SUPPORTING INFORMATION

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