Rates of opioid misuse, abuse, and addiction in chronic pain: a systematic review and data synthesis

Kevin E. Vowles, Mindy L. McEntee, Peter Siyahhan Julnes, Tessa Frohe, John P. Ney, David N. van der Goes

Abstract

Opioid use in chronic pain treatment is complex, as patients may derive both benefit and harm. Identification of individuals currently using opioids in a problematic way is important given the substantial recent increases in prescription rates and consequent increases in morbidity and mortality. The present review provides updated and expanded information regarding rates of problematic opioid use in chronic pain. Because previous reviews have indicated substantial variability in this literature, several steps were taken to enhance precision and utility. First, problematic use was coded using explicitly defined terms, referring to different patterns of use (ie, misuse, abuse, and addiction). Second, average prevalence rates were calculated and weighted by sample size and study quality. Third, the influence of differences in study methodology was examined. In total, data from 38 studies were included. Rates of problematic use were quite broad, ranging from <1% to 81% across studies. Across most calculations, rates of misuse averaged between 21% and 29% (range, 95% confidence interval [CI]: 13%-38%). Rates of addiction averaged between 8% and 12% (range, 95% CI: 3%-17%). Abuse was reported in only a single study. Only 1 difference emerged when study methods were examined, where rates of addiction were lower in studies that identified prevalence assessment as a primary, rather than secondary, objective. Although significant variability remains in this literature, this review provides guidance regarding possible average rates of opioid misuse and addiction and also highlights areas in need of further clarification.

Keywords: Opioids, Chronic pain, Problematic use, Abuse, Addiction, Misuse

1. Introduction

In the treatment of chronic pain, there may be no area of greater controversy than the use of opioids. Changes in attitudes with respect to opioid use toward the end of the 20th century, and subsequent exponential increases in use, have been well documented. More recently, the burgeoning public health issue regarding opioid-related adverse events has perhaps been equally well documented, as the use of opioids in chronic pain brings with it marked potential for adverse events for the patient, including overdose, experience of physiological dependence and subsequent withdrawal, addiction, and negative impacts on functioning. Attention paid to the so-called “opioid epidemic” (eg, Refs. 19,32) has highlighted the need to clearly differentiate and identify the types of problematic prescription opioid use (eg, misuse, abuse, addiction) and discern their frequency in treated patients with chronic pain.

Attempts to calculate rates of problematic opioid use behavior have suffered from imprecise and poorly defined terminology. Two recent sets of expert consensus statements, one suggesting a framework for measuring abuse liability for use in trials of analgesics for those with chronic pain and the other a set of definitions for opioid-related adverse events, identified 8 loose and overlapping categories of problematic use, including misuse, abuse, addiction, aberrant use, dependence, nonmedical or nontherapeutic use, physical dependence, and psychological dependence (also see the review of Webster and Fine, who further define “pseudoaddiction”). The vagueness inherent in these definitions, areas of overlap among them, and their sometimes interchangeable use have made it difficult to determine exact rates and types of problematic opioid use. For example, the narrative review of Højsted and Spøgren detailed the findings of 25 studies involving patients with chronic pain prescribed with opioids, concluding that the prevalence of problematic opioid use behavior ranged from 0% to 50%. Although this span was representative of the literature at the time, it was of questionable value for delineating the scope, impact, and prevalence of the problem or in facilitating informed clinical and policy decisions regarding the allocation of screening and treatment resources. Martell et al., in their review of opioid use for low back pain, reported a similar range of current problematic opioid use (3% to 43%).

The purpose of this study was to perform an updated review of problematic opioid use in chronic pain using explicitly defined terms for rates of problematic use in the literature. We synthesized the data to clarify and calculate prevalence estimates to increase precision and utility. As a secondary set of analyses, we investigated whether variation in the rates of problematic opioid use were related to differences in study characteristics (ie, primary study purpose, study design, method of assessment, clinical setting).

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

# Department of Psychology, University of New Mexico, Albuquerque, NM, USA
# Department of Neurology, University of Washington, Seattle, WA, USA
# Department of Economics, University of New Mexico, Albuquerque, NM, USA
*Corresponding author. Address: Department of Psychology, University of New Mexico, MSC03 2220, Logan Hall, 1 University of New Mexico, Albuquerque, NM 87131, USA. Tel: +1 505-277-4121; fax: +1 505-277-1394. E-mail address: k.e.vowles@gmail.com (K. E. Vowles).

PAIN 156 (2015) 569–576
© 2015 International Association for the Study of Pain
http://dx.doi.org/10.1097/j.pain.0000460357.01998.f1
2. Methods

2.1. Literature search strategy

We searched the clinical and scientific literature using Science Direct, Google Scholar, PubMed, and PsychINFO/PsycArticles databases for articles published between January 2000 and January 2013. We repeated the search in November 2013 to include articles published or accepted since January 2013. We used broad search terms to increase the probability of accurate identification of target articles (Table 1). We also reviewed reference lists to identify any articles that the initial search had missed.

2.2. Abstract screening

The abstracts of all studies identified in the literature search were read by 2 reviewers to assess eligibility for full-text data extraction. To be eligible for data extraction, studies met the following criteria: (1) only adult participants (ie, 18+ years of age), (2) sample composed of individuals with chronic noncancer pain (persistent pain lasting longer than 3 months), (3) participants were using opioids orally (to exclude studies of opioids delivered transdermally or through injection/intrathecal pump), (4) the abstract listed 1 or more of the following terms in reference to opioid use: abuse, misuse, dependence, addiction, or aberrant/problematic behavior, and (5) quantitative information was provided (as opposed to a commentary or qualitative review) regarding rates of problematic opioid use.

2.3. Full-text data extraction

Each study fitting the inclusion criteria was read in full by 2 members of the study team to extract and record data on a standardized data extraction form. The extracted information included participant demographics and pain details (ie, sample size, gender, age, pain duration, ethnicity, pain location), primary objective (eg, assessment of prevalence, medication safety/efficacy), study design (ie, cross-sectional/prospective/retrospective), study setting details, country of data collection, and method used to identify problematic opioid use (ie, structured/unstructured clinical interview, urine drug screen [UDS], chart review, clinical judgment, questionnaire).

2.3.1. Coding of current opioid misuse, abuse, and addiction

Problematic use of opioids was categorized according to recent consensus statements published by the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT)44 and Analgesic, Anesthetic, and Addiction Clinical Trials, Translations, Innovations, Opportunities, and Networks (ACTION)53 panels, an overlapping group of experts with representation from private, public, and governmental domains. Based on these consensus statements, and the associated commentaries of Butler and Sullivan,54 the following definitions were used to categorize problematic use as misuse, abuse, or addiction.

(1) Misuse: Opioid use contrary to the directed or prescribed pattern of use, regardless of the presence or absence of harm or adverse effects.

(2) Abuse: Intentional use of the opioid for a nonmedical purpose, such as euphoria or altering one’s state of consciousness.

(3) Addiction: Pattern of continued use with experience of, or demonstrated potential for, harm (eg, “impaired control over drug use, compulsive use, continued use despite harm, and craving”).53

2.3.2. Rating of study quality

The quality of each study was rated using 8 of the 9 criteria used by Chou et al.,21 which determines whether the study evaluated test performance in a population other than the one used to derive the instrument (ie, derivation vs validation study), was coded but eventually discarded as it was deemed less useful in discriminating between high and low quality. The following guidelines were used to code the type of problematic use. First, the percentage of study participants meeting criteria for each type of problematic use was extracted and recorded, where possible, to the tenths decimal place. A single percentage was recorded from studies that met the criteria for only a single type of problematic use, whereas studies that reported separately on more than 1 type of problematic use provided more than 1 estimate (eg, 1 for participants meeting criteria for misuse and 1 for participants meeting criteria for addiction). Second, when studies reported a range of values regarding the percentage of patients meeting criteria for 1 type of problematic use, a minimum and maximum value was recorded. Third, only current problematic opioid use was recorded; data were not used if a study reported only on historical or lifetime problematic opioid use. When insufficient or ambiguous information was provided in the published articles or available supplemental data, we contacted study authors for additional details.

2.3.1. Coding of current opioid misuse, abuse, and addiction

Problematic use of opioids was categorized according to recent consensus statements published by the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT)44 and Analgesic, Anesthetic, and Addiction Clinical Trials, Translations, Innovations, Opportunities, and Networks (ACTION)53 panels, an overlapping group of experts with representation from private, public, and governmental domains. Based on these consensus statements, and the associated commentaries of Butler and Sullivan,54 the following definitions were used to categorize problematic use as misuse, abuse, or addiction.

(1) Misuse: Opioid use contrary to the directed or prescribed pattern of use, regardless of the presence or absence of harm or adverse effects.

(2) Abuse: Intentional use of the opioid for a nonmedical purpose, such as euphoria or altering one’s state of consciousness.

(3) Addiction: Pattern of continued use with experience of, or demonstrated potential for, harm (eg, “impaired control over drug use, compulsive use, continued use despite harm, and craving”).53

2.3.2. Rating of study quality

The quality of each study was rated using 8 of the 9 criteria used by Chou et al.,21 which determines whether the study evaluated test performance in a population other than the one used to derive the instrument (ie, derivation vs validation study), was coded but eventually discarded as it was deemed less useful in discriminating between high and low quality. The remaining 8 criteria evaluate study sampling issues (eg, consecutive sample or random subset, proportion of missing
data), adequate description of study methods (eg, sample and patterns of opioid prescription, criteria to identify problematic behavior), and potential influence of raters on identification of problematic behavior (eg, rater blinding regarding the identification of problematic use). Consistent with Chou et al., studies that met the majority of the criteria (5 or more) were regarded as higher quality.

2.4. Analytic plan

Extracted data were entered into SPSS (version 21; IBM Corporation). The primary variables of interest were average rates of misuse, abuse, and addiction across studies. Because a small number of studies reported these rates as a range of values, 2 sets of calculations were performed for each type of problematic use, a minimum and a maximum. When only a single value was recorded, that value was entered as both the minimum and maximum value as that ensured that both the minimum and maximum calculations included the complete set data. Although we expected minimum and maximum values to be close to one another, this method of calculation was deemed to make best use of all available data and allow equal weightings for each study’s data.

The first analytic step involved the calculation of unweighted raw means and SDs for rates of misuse, abuse, and addiction across all included studies. In addition, we calculated a number of weighted means, including weighting for raw sample size and log-transformed sample size. The log transformation was performed to address the large variability in sample size and apparent exponentiation of the sample size distribution within the largest studies. In addition, a Winsorizing procedure was performed for studies with sample sizes of greater than 1334 participants, which was the point at which outliers were identified within stem-and-leaf plots; there was also evidence of a bimodal distribution at this cut-point. For the analyses using the Winsorized sample size data, samples size for all studies with greater than 1334 participants were set to 1334, the value of the next largest sample size.

In addition to the analyses involving weightings by sample size, weighted means were calculated for study quality. Furthermore, means for studies of high and low quality were evaluated separately. Finally, a weighted interaction term of log-transformed sample size and quality rating was calculated using standardized scores (z-scores).

As a secondary set of analyses, differences in rates of problematic use were assessed in relation to primary study purpose (ie, Was the assessment of prevalence of misuse, abuse, and addiction the primary aim?), study design (ie, retrospective, cross-sectional, prospective), method of identification (eg, questionnaire, structured/semistructured interview, chart review, UDS), and clinical setting (eg, primary care, pain clinic). A series of analyses of variance (ANOVAs) was used to analyze for differences in rates of problematic use based on these study characteristics.

3. Results

3.1. Search results

Figure 1 displays the flow of information throughout the different phases of the search in a manner consistent with the PRISMA statement. The search yielded a total of 311 records for screening after the exclusion of 46 nonempirical papers, such as reviews, letters, and commentaries. An additional 9 records were identified in the updated, November 2013, search, yielding a total of 320 records for screening.

Kappa values indicated an acceptable level of agreement among raters, range $\kappa = 0.79$ to 0.91. All articles that had a discrepant rating after this stage of evaluation were retained for full-text data review.

A total of 78 articles were retained for full-text review. Of these, 40 were excluded for the reasons outlined in Figure 1. Therefore, 38 articles were used in data synthesis.

3.2. Characteristics of selected studies

Individual study characteristics are located in Table 2. The majority of studies, 35 (92%), reported on either misuse or addiction, whereas the remaining 3 studies reported on both. In total, 29 studies (76%) reported on rates of misuse and 12 (32%) on rates of addiction. Abuse was reported in only a single study, that of Banta-Green et al., as this was the only study that reported specifically on participant intention. Therefore, no further calculations of abuse prevalence were performed.

Generally, considerable variability regarding study characteristics was apparent. Sample size, for example, ranged from 63 to 938,586 participants. Quality ratings ranged from 0 to 8. Sample size and quality ratings were significantly and negatively correlated with one another, $r = -0.36, P < 0.05$. There was also variability in reporting basic demographic and pain-related information. Specifically, 77.5% of studies reported on participant sex, 70.0% provided some information on age (with 15.0% providing nonnumeric information that could not be extracted—eg, "most patients fell into the 35- to 50-year-old range"). 47.5% provided information of any kind on participant ethnicity, and only 22.5% provided information on education. Regarding pain-related information, a minority of studies provided information on pain location (42.5%), or information on pain duration (37.5%).

---

**Figure 1**. Flow of information thought the different phases of the review, as specified by the PRISMA statement.
### 3.3. Rates of opioid misuse and addiction

Overall, sizeable variability in rates of both misuse and addiction was indicated across reviewed studies. Rates of abuse ranged from 0.08% to 81.0% and addiction rates ranged from 0.7% to 34.1% across all studies. For high-quality studies (n = 13 for misuse and 10 for addiction), misuse rates ranged from 2.0% to 56.3% and addiction rates from 0.7% to 23.0%. For low-quality studies (n = 16 for misuse and 2 for addiction), misuse rates...
Table 3
Opioid misuse—unweighted and weighted means, SDs, and 95% confidence interval (CI).

<table>
<thead>
<tr>
<th>Quality</th>
<th>Unweighted</th>
<th>Weighted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum, %</td>
<td>Maximum, %</td>
</tr>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>95% CI</td>
</tr>
<tr>
<td>Low-quality</td>
<td>23.6 (16.4)</td>
<td>14.7-32.5</td>
</tr>
<tr>
<td>High-quality</td>
<td>31.8 (31.2)</td>
<td>16.5-47.1</td>
</tr>
</tbody>
</table>

*Interaction term the product of standardized scores for the log transformed sample size and quality rating.

Table 4
Opioid addiction—unweighted and weighted means, SDs, and 95% confidence interval (CI).

<table>
<thead>
<tr>
<th>Quality</th>
<th>Unweighted</th>
<th>Weighted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum, %</td>
<td>Maximum, %</td>
</tr>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>95% CI</td>
</tr>
<tr>
<td>Low-quality</td>
<td>8.8 (7.3)</td>
<td>4.3-13.3</td>
</tr>
<tr>
<td>High-quality</td>
<td>23.1 (12.9)</td>
<td>3.4-39.2</td>
</tr>
</tbody>
</table>

*Interaction term the product of standardized scores for the log transformed sample size and quality rating.

3.4. Comparisons of study design, diagnostic method, and clinical setting

As noted, because the studies identified for data extraction were quite varied regarding their characteristics, we examined rates of misuse and addiction across studies regarding primary study purpose, study design, assessment method used to identify problematic behavior, and clinical setting. For each of these 4 variables, 4 ANOVAs were conducted (minimum/maximum; misuse/addiction). A Bonferroni correction was used for all pairwise comparisons to help control against the commission of a type I error.

Across all analyses, results indicated only 1 significant difference in relation to study characteristics. Specifically, mean unweighted rates of opioid addiction were lower in the 7 studies that identified the assessment of prevalence as the primary study objective, minimum/maximum mean = 5.5%/6.2% (SD = 4.6%/6.2%; 95% CI = 2.1%-10.8%), in comparison with 5 studies for which prevalence assessment was a secondary objective, minimum/maximum mean = 18.4%/19.4% (SD = 10.7%/9.4%; 95% CI = 9.0%-27.6%), all F > 8.3, all P < 0.02.

For opioid misuse, 11 studies (37.9%) identified the assessment of prevalence as the primary study aim and 18 studies (62.1%) as a secondary aim. No significant differences were indicated in average rate of misuse across studies, all F < 2.7, all P > 0.11.

All other comparisons did not indicate any significant differences in relation to the additional study characteristics evaluated. Specific findings are detailed in the following paragraphs and descriptive information is provided in Table 5.

Regarding study design, of the 38 studies reviewed, 39.5% were prospective, 34.2% were cross-sectional, and 26.3% were retrospective. No significant differences were indicated by any of the analyses comparing rates of misuse and addiction with design, all F < 1.0, all P > 0.37.

The assessment method used also varied substantially across studies with the majority, 64.9%, using only a single assessment method (questionnaire: 21.6%, clinical judgment/chart review: 21.6%; structured/semistructured interview: 13.5%; UDS: 8.1%). The remaining 35.1% of studies used a UDS plus at least 1 other method, which were coded as a single assessment category.
Accurate identification and enumeration of problematic opioid use in those with chronic pain is important. Our review evaluated the current state of the literature regarding rates of opioid misuse, abuse, and addiction in chronic pain. The results are concordant with previous work in many ways. Chiefly, the substantial variability in studies evaluating problematic opioid use remains apparent as there were many designs used, methods of identification used, and study settings examined. The range of rates of problematic use was even broader than that has been previously reported with weighted means and also have the added benefit of allowing for greater precision in the language used in relation to patterns of opioid use in chronic pain. As these analyses were likely underpowered because small cell sizes and the ranges analyzed were broad, these results ought to be interpreted cautiously, and we include them here to primarily provide information of potential use to future studies in this area.

We can make several recommendations for future studies of problematic opioid use in chronic pain. First, studies must specify the relevant demographic and pain-related details. At a minimum, we suggest that these include gender, age, and ethnicity, as well as pain location and duration. These details were included in a surprisingly small number of studies despite their demonstrated relevance in treatment response and role in the potential for problematic opioid use. The inclusion of measures of pain intensity and interference would likely provide valuable additional information. Second, there is likely a benefit to be found in specifying type of problematic use that is being assessed and specifically designing studies to evaluate prevalence as a primary objective. Such specification may aid in decreasing variability across studies regarding rates of problematic use and perhaps also have the added benefit of allowing for greater precision in the language used in relation to patterns of opioid use in chronic pain. Third, at present, there is no clear gold standard for use in the identification of misuse, abuse, and addiction. Perhaps the most thorough method is the Aberrant Drug Behavior Index (ADBI) used by Butler et al. The ADBI involves a triangulation approach consisting of self-reported patterns of opioid use evaluated by a structured interview, physician-reported patterns of use, and a UDS. A positive ADBI, indicating the presence of problematic opioid use, consisted of either a positive rating on the structured clinical interview or positive ratings on both the physician report and UDS. In this review, this triangulation method was coded as indicating misuse, but it seems feasible to modify it so that it also provides information regarding abuse and addiction.

The results of this review have 2 key implications. First, misuse and addiction do seem to be distinct patterns of problematic opioid use, at least based on the definitions used here and the differences in observed mean rates between them. Second, misuse seems more common than addiction. Several types of misuse were identified within studies and included underuse, erratic or disorganized use, inappropriate use (eg, to manage symptoms of anxiety or other sorts of distress), use in conjunction with alcohol or illegal substances (eg, marijuana), and, of course, overdose. If it is accurate that approximately 1 in 4 patients on opioids display patterns of opioid misuse, but not addiction, then perhaps more efficient targeting of treatment resources would be of benefit. Some forms of misuse, for example, may be readily addressed through relatively low-intensity methods such as education or frequent follow-up visits. One prominent example of a fairly low-intensity intervention is that of Jamison et al., who held monthly meetings with patients deemed to be at “high risk” of opioid misuse. These meetings were a combination of motivational approaches, opioid education, and opioid use monitoring, including a UDS, held monthly over the course of 6 months. At the conclusion of the study period, the documented
rates of aberrant behavior was low and comparable to rates documented for another group of patients, who were deemed to be of "low risk" of opioid misuse at the onset of the study. These findings suggest that there are alternatives available to providers who treat high-risk patients beyond simply not prescribing the medications at all. A more recent study from the same group further highlights a potential key role of cravings in opioid misuse, which presents another option for intervention given that the substance abuse literature already provides effective interventions directed at altering the impact of drug cravings more generally, and these could perhaps be readily adapted to problematic opioid use.\textsuperscript{7,63,64}

The results of this review have several limitations. The most obvious is the degree of variability within this literature. In spite of our attempts to minimize the impact of this variability, the range of misuse and addiction was incredibily broad, as were measures of dispersion. Furthermore, there are other potential sources of variability in findings that were not possible to code and extract in a uniform manner. These include duration of opioid use, history of nonopioid substance misuse, abuse, or addiction, dosage levels and frequency of use, as well as health care system variables, such as frequency of prescription reviews, drug testing, or use of opioid "contracts." These sources of variability will likely continue to cloud our ability to make precise estimates. There is clearly room here for a series of carefully controlled studies where sources of variability are held constant, or as constant as possible, to more clearly illuminate prevalence rates of problematic opioid use in individuals with chronic pain.

There was 1 curious finding that we have not yet emphasized. The overwhelming majority of studies within this review took place in the United States. Only 3 of the 38 studies took place in other countries, which suggests that this issue is of both high interest and is perhaps a problem that is somehow uniquely relevant to the US. The latter interpretation is supported by the finding of Manchikanti et al.\textsuperscript{31} indicating that the US population, which represents approximately 5% of the Earth's population, consumed approximately 80% of the global supply of prescribed opioids in the first decade of this century. This is an intriguing issue and although there are likely many factors involved, neither the abundance of opioids prescribed for the treatment of chronic pain nor the large proportion of studies of problematic opioid use seem to have helped diminish the prevalence, impact, or cost of chronic pain in the US since the explosion in opioid use for chronic pain.\textsuperscript{22}

One final, related, comment on the use of opioids in chronic pain seems appropriate. In short, it is not clear whether the risks of opioid use outweigh the potential for benefit. The efficacy of opioids and their suitability for the long-term management of chronic pain still remain very much in question\textsuperscript{3,4,13,51,54,55} and while this uncertainty in effectiveness is well established, it stands in somewhat stark contrast to the clinical reality of chronic pain treatment, where rates of prescriptions have skyrocketed such that opioids are now among the most frequently prescribed medications. What does seem clear, however, is that the rapid increase in opioid use has had what Sullivani\textsuperscript{34} referred to as "unintended" consequences that, for at least some patients, require an additional form of intervention to curtail patterns of problematic use and potential for harm. We are not certain whether the benefits derived from opioids, which are rather unclear based on the extant literature, compensate for this additional burden to patients and health care systems.

Conflict of interest statement
The authors have no conflicts of interest to declare.

This research was supported by a grant from the Center for Health Policy at the Robert Wood Johnson Foundation to the first and last authors.

Acknowledgements
The authors thank Dr Robert Valdez for his insightful and helpful comments on a previous version of this report.

Article history:
Received 13 May 2014
Received in revised form 27 December 2014
Accepted 29 December 2014

References
\begin{enumerate}
\item Butler S. The IMMPACT factor or IMMPACT strikes again! Pain 2013;154:2243–4.
\item Chełminski PR, Ives TJ, Felix KM, Prakken SD, Miller TM, Perhac JS, Malone RM, Bryant ME, DelWalt DA, Pignone MP. A primary care, multi-disciplinary disease management program for opioid-treated patients with chronic non-cancer pain and a high burden of psychiatric comorbidity. BMC Health Serv Res 2005;5:3.
\end{enumerate}


