A literature review of the impact of exclusion criteria on generalizability of clinical trial findings to patients with chronic pain

Vafi Salmasia,*, Theresa R. Liia, Keith Humphreysb,c, Vinay Reddya, Sean C. Mackeya

Abstract
The ability of clinical trials to inform the care of chronic pain may be limited if only an unrepresentative subset of patients are allowed to enroll. We summarize and report new insights on published studies that report on how trial exclusions affect the generalizability of their results. We conducted a PubMed search on the following terms: ("eligibility criteria" AND generalizability) OR ("exclusion criteria" AND generalizability) OR “exclusion criteria” OR “eligibility criteria” AND pain. We only considered studies relevant if they analyzed data on (1) the prevalence and nature of exclusion criteria or (2) the impact of exclusion criteria on sample representativeness or study results. The 4 articles that were identified reported differences in patients who were included and excluded in different clinical trials: excluded patients were older, less likely to have a paid job, had more functional limitations at baseline, and used strong opioids more often. The clinical significance of these differences remains unclear. The pain medicine literature has very few published studies on the prevalence and impact of exclusion criteria, and the outcomes of excluded patients are rarely tracked. The frequent use of psychosocial exclusions is especially compromising to generalizability because chronic pain commonly co-occurs with psychiatric comorbidities. Inclusion of more representative patients in research samples can reduce recruitment barriers and broaden the generalizability of findings in patients with chronic pain. We also call for more studies that examine the use of exclusion criteria in chronic pain trials to better understand their implications.

Keywords: Eligibility criteria, Exclusion criteria, Generalizability, Pain, Psychological

1. Introduction
Roughly 20% of adults in the United States suffer from chronic pain, with 8% experiencing high-impact chronic pain; this costs the economy more than $600 billion annually. Patients with chronic pain report more medical comorbidities than patients without chronic pain. Chronic pain also co-occurs frequently with psychiatric disorders such as depression and anxiety. Yet, randomized controlled trials (RCTs) for chronic pain disorders often exclude patients with comorbidities in the pursuit of participant safety, reducing treatment response heterogeneity, and increasing statistical efficiency. This exclusionary approach, while often necessary for efficacy or mechanistic studies, comes at the cost of having study samples that poorly represent the real-world population with the studied disease. The nature of the research question usually determines if the eligibility criteria need to be more exclusive vs more inclusive. Explanatory trials determine efficacy of an intervention under ideal circumstances and thus requires a more homogeneous population. However, these trials have to be followed by more pragmatic trials to determine the effectiveness of the same intervention under general conditions that it is intended to be applied.

The degree to which the results from a study can be applied to a larger or more heterogeneous population is referred to as generalizability, or external validity. In recent years, the generalizability of RCTs has come under scrutiny in several disciplines, including psychiatry, rheumatology, and infectious
disease. Whether this same problem occurs in pain medicine has not been studied. We therefore reviewed available studies that evaluate differences in baseline characteristics or outcomes between included and excluded participants in chronic pain trials, and we summarize the impact of these exclusions on study results when applicable.

2. Methods

We closely emulated the methods of the Cross-disease Review of Exclusion Across Medicine (CREAM) project previously presented by Humphreys and colleagues. Their project assessed the impact of exclusion criteria in research conducted across a range of medical specialties. A detailed description of the literature review procedure can be found in the article on schizophrenia trials by Humphreys. We applied the same search strategy and terms used by the CREAM project, but with a different condition of interest (pain). We did not preregister our protocol because our methods more closely resemble a literature review, rather than a systematic review.

Literature was primarily identified by conducting English-language searches in PubMed (original date of search: November 14, 2020) on the following terms: "Eligibility criteria and generalizability" (anywhere in article), "exclusion criteria and generalizability" (anywhere in article), "exclusion criteria" (in the title of article) and "eligibility criteria" (in the title of article) (ie. "(eligibility criteria" AND generalizability) OR (exclusion criteria" AND generalizability) OR "exclusion criteria"[ti] OR "eligibility criteria"[ti] AND pain). This generated 822 articles. Like other CREAM projects, we then focused on patients with chronic pain by adding the search term "pain (anywhere in the article)" to the other search terms; this resulted in 24 articles. We considered studies relevant if they analyzed data on (1) the prevalence and nature of exclusion criteria and/or (2) the impact of exclusion criteria on sample representativeness or study results. We excluded studies that only reported their exclusion criteria and rate. The first authors (V.S. and T.L.) independently reviewed all these articles and their references to other studies of exclusion criteria (Fig. 1). Both authors chose similar studies that met our relevance criteria.

3. Results

One clinical trial and 3 review articles (total N = 4) met our relevance criteria. The clinical trial evaluated differences in baseline characteristics and outcomes between included and excluded participants. The remaining articles reported on the prevalence and nature of exclusion criteria. Disease entities covered include low back pain (2 articles), knee osteoarthritis (1 article), and postoperative pain after hip and knee arthroplasty (1 article).

Maas et al. analyzed the outcomes of 1148 participants in a pragmatic multicenter trial to treat low back pain. They compared the outcomes of a randomized group (n = 241) to 2 observational groups who received the studied treatment without randomization. Observational group 1 (n = 455) consisted of participants who fulfilled all eligibility criteria but were not randomized. Observational group 2 (n = 452) consisted of participants who did not meet eligibility criteria. Both observational groups 1 and 2 received the study treatment. Their sample was drawn from 16 Dutch clinics and 102 physiotherapy practices. The authors explain that the eligibility criteria of the randomized group were "standardized" and they were similar to the ones used by pain physicians in clinical practice in the Netherlands.

The most frequently applied exclusion was "psychological problems" which prevented 60.8% of participants in observational group 2 from enrolling in the RCT, including those with other exclusions (17.5%). Age and body mass index were the second (31.2%) and third (16.2%) most frequently applied exclusions, respectively. Participants with multiple overlapping criteria made up 19.2% of excluded participants. The combined effect of all criteria was to exclude 452 of 1148 (39.4%) participants who had otherwise been deemed eligible during initial screenings but were instead diverted to observational group 2. Compared to participants in the randomized group and observational group 1, those in observational group 2 were older, less likely to have a paid job, had more functional limitations at baseline, and used strong opioids more often.

Regarding outcomes from the study by Maas et al., patients who had pain from facet joints had slightly less pain relief if they were in observational group 2 compared with randomized patients (mean difference of 0.4 on a numeric scale of 1–10). No differences in pain scores were found in observational group 2 patients who had pain from sacroiliac (SI) joints. Patients in observational group 2 also had slightly less improvement in functional status (mean difference of 5.83 and 7.16 on the Oswestry Disability Index [pain from facet joints and SI joints, respectively]), and a lesser likelihood of perceiving treatment success (odds ratio 0.67 and 0.72 [pain from facet joints and SI joints, respectively]). The authors report that these statistically significant differences were too small to be clinically relevant.

Koog et al. reviewed 355 RCTs for knee osteoarthritis which were published up to December 2011. They divided exclusion criteria into “non–osteoarthritis-related” and “osteoarthritis-related” reasons. Ten exclusion criteria were identified in the non–osteoarthritis-related category. Inability to give informed consent was the most commonly applied non–osteoarthritis-related exclusion (used in 79.4% of trials), followed by recent exposure to other potential treatments (61.4%) and medical conditions that make it difficult to participate in a trial (57.2%) or might pose a risk to receiving treatment (56.0%). Other psychological or psychiatric exclusions were not identified in the review; however, it is unclear whether these were included under “medical conditions.” In the osteoarthritis-related category, 9 overlapping exclusion criteria were identified. Lower boundaries of age were the most common osteoarthritis-related exclusion (58.9%), followed by exclusion of nonpainful osteoarthritis (58.3%) and limits on osteoarthritis severity (47.3%). Koog et al. did not tabulate the number of included or excluded participants, nor did they compare baseline characteristics or outcomes of excluded participants to included participants.

Amundsen et al. reviewed 168 RCTs for chronic low back pain published between 2006 and 2012. Exclusion criteria were noted to be highly diverse, so the authors categorized exclusion criteria into 19 themes. The most ubiquitous exclusion themes were causes of back pain from other conditions, including malignancy (46.4% of trials) or trauma (37.5%). Pregnancy was a common exclusion criterion (48.2%) as was having previous or scheduled surgery (50%). Note worthy exclusion criteria frequently encountered in real-world patients include psychosocial conditions (ie, psychiatric disorders, depression, severe psychiatric disorder, and impaired cognition) (34.5%) and medicolegal issues (23.2%). Amundsen et al. did not tabulate the number of included or excluded participants, nor did they compare baseline characteristics or outcomes of excluded participants to included participants.

Finally, Pedersen et al. analyzed the exclusion criteria used in 550 RCTs investigating postoperative pain treatment for hip and
knee total arthroplasty. The authors also compared baseline characteristics from 48,962 research participants enrolled in RCTs with an unrandomized clinical cohort from the Danish Hip and Knee Arthroplasty Registries from 2005 to 2019. The authors found that the most commonly used exclusion criteria were having a high American Society of Anesthesiologists (ASA) Physical Status Classification score (62% of trials), older age (45%), chronic opioid use (41%), neurological disorders (41%), and renal disease (40%). Psychiatric or cognitive exclusions were used in 29% of trials.18

Research participants were on average younger than their clinical counterparts; however, body mass index (BMI) American Society of Anesthesiologists (ASA) physical status scores, and sex distribution were comparable. When examining trends over time, the enrolled research participants’ average age has increased, and in the subgroup of patients undergoing total knee arthroplasty, BMI and ASA scores have also increased. Meanwhile, age, BMI, and ASA scores in the research cohort have not changed, although fewer women are being included in total hip arthroplasty trials over time. The prevalence of other exclusion criteria (such as chronic opioid use or medical and psychiatric comorbidities) was not reported in either cohort. The total number of included or excluded participants from the research cohort was not reported either.18

4. Discussion

The current review found very few published studies on the prevalence and impact of exclusion criteria in clinical trials for chronic pain. Disease entities were limited to knee osteoarthritis, pain after hip and knee arthroplasty, and chronic low back pain.1,11,13,18 This leaves out large swaths of other common chronic pain conditions such as migraine, lumbar radiculopathy, and peripheral neuropathy. One study compared the outcomes of randomized patients with excluded patients who had received the studied treatment outside of the RCT.13 Three review articles reported on the nature and prevalence of exclusion criteria used in clinical trials1,11,16, only one of them analyzed differences in baseline characteristics by comparing trial patients with a national database of nontrial patients.18 None of the review articles estimated exclusion rates, probably because the number of patients excluded before randomization is inconsistently reported.1,11,18

The authors of the study that followed excluded patients with chronic low back pain argued that the differences in outcomes between randomized and excluded patients were too small to be clinically relevant, despite reaching statistical significance. We agree that the mean differences seen with numeric pain scores and Oswestry Disability Index scores are relatively inconsequential.13 Exclusions are usually designed assuming that they result in more heterogeneous outcomes. The interpretation of the authors shows that this assumption is wrong and more strict exclusion criteria were unnecessary. However, we believe that the difference in perceived treatment success is clinically impactful because patients who believe that their treatment was unsuccessful are more likely to seek additional medical care and less likely to return to work, with resultant costs to the healthcare system and society.

Another way to interpret the difference in perceived treatment success is by calculating the number needed to harm (NNH) for having exclusionary comorbidities. Number needed to harm measures how many people need to be exposed to a risk factor for 1 person to experience an adverse effect or harm. A lower NNH indicates a greater risk of harm. In this case, if we define “harm” as a lack of perceived treatment success 3 months after receiving treatment, we can use the formula $\text{NNH} = 1/(\text{IE} - \text{I}_0)$, where $\text{I}_E$ is the incidence of harm in those with exclusionary comorbidities and $\text{I}_0$ is the incidence of harm in those without exclusionary comorbidities. Based on data reported by Maas et al., the NNH for having exclusionary comorbidities was 8 for...
patients with pain from facet joints and 12 for patients with pain from SI joints. This means that for every 8 patients with exclusionary comorbidities receiving treatment for facet joint pain, 1 patient will report a lack of treatment success 3 months after treatment. Likewise, for every 12 patients with exclusionary comorbidities receiving treatment for SI joint pain, 1 patient will report a lack of treatment success 3 months after treatment. As a point of reference, the NNH for gabapentin causing somnolence² is reported to be 11.

One class of exclusions that stood out in our review was psychiatric and psychosocial problems. Approximately one-third of all low back pain trials excluded participants with psychological issues. This proportion was slightly less in trials for pain after hip and knee arthroplasty. Chronic pain is frequently comorbid with psychiatric disorders such as depression and anxiety, and the presence of adverse psychosocial factors portends a worse prognosis in both surgical⁶,23,25 and nonsurgical⁶,12,16 populations. By excluding these patients, the researchers hope to reduce heterogeneity in treatment response; however, this practice likely results in overestimating treatment efficacy in a general chronic pain population. Some clinicians may reserve procedural interventions for patients without psychiatric disease, making their clinical population more comparable to research populations. However, this still leaves a broad category of patients with chronic pain. Ultimately, such studies are important to determine whether the inclusion of more representative patients in research samples can reduce recruitment barriers and broaden the generalizability of study findings across chronic pain populations.

We also recommend more transparent reporting of exclusion criteria in chronic pain trials because this will promote interpretability of results and facilitate future research on the effects of exclusion criteria. Finally, we call for more inclusive studies with less stringent exclusion criteria because this will promote generalizability of results to real-world patient populations. Although increased heterogeneity in the study population will necessitate larger sample sizes to discern an effect of an intervention, pragmatic trial designs embedded into existing practice settings can be used to recruit large numbers of participants at lower cost than conventional trials.¹⁹

Disclosures

The authors have no conflict of interest to declare.

Article history:

Received 4 May 2022
Received in revised form 12 September 2022
Accepted 4 October 2022

References

[1] Amundsen PA, Evans DW, Rajendran D, Bright P, Bjork T, Eldridge S, Buchbinder R, Underwood M, Froud R. Inclusion and exclusion criteria used in non-specific low back pain trials: a review of randomised controlled trials published between 2006 and 2012. BMC Musculoskelet Disord 2018;19:113.
[2] Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity: a literature review. Arch Intern Med 2003;163:2433–45.
[3] Dahlgren J, Lucas J, Zelaya C, Nahin R, Mackey S, DeBar L, Kems R, Von Korff M, Porter L, Helmc G. Prevalence of chronic pain and high-impact chronic pain among adults—United States, 2016. MMWR Morb Mortal Wkly Rep 2018;67:1001–6.
[4] Dominick CH, Blyth FM, Nicholas MK. Unpacking the burden: understanding the relationships between chronic pain and comorbidity in the general population. PAIN 2012;153:293–304.
[5] Edwards RR, Dworkin RH, Sullivan MD, Turk DC, Wasan AD. The role of pain comorbid processes in the development and maintenance of chronic pain. J Pain 2016;17:T70–92.
[6] Gandhi M, Ameli N, Bacchetti P, Sharp GB, French AL, Young M, Gange SJ, Anastos K, Holman S, Levine A, Greenblatt RM. Eligibility criteria for HIV clinical trials and generalizability of results: the gap between published reports and study protocols. AIDS 2005;19:1885–96.
[7] Garcia D. Gabapentin for chronic neuropathic pain. Am Fam Physician 2015;92:Online.
[8] Gheonne MM, O’Hara MW. Depression and postoperative complications: an overview. BMC Surg 2016;16:5.
[9] Goldstein-Piekarski AN, Williams LM, Humphreys K. A trans-diagnostic review of anxiety disorder comorbidity and the impact of multiple exclusion criteria on studying clinical outcomes in anxiety disorders. Transl Psychiatry 2016;6:e847.
[10] Humphreys K. A review of the impact of exclusion criteria on the generalizability of schizophrenia treatment research. Clin Schizophr Relat Psychoses 2017;11:49–57.
[11] Koog YH, Wu H, Jung WY. Eligibility criteria in knee osteoarthritis clinical trials: systematic review. Clin Rheumatol 2013;32:1569–74.
[12] Kroenke K, Wu J, Bair MJ, Krebs EE, Damush TM, Tu W. Reciprocal relationship between pain and depression: a 12-month longitudinal analysis in primary care. J Pain 2011;12:964–73.
[13] Maas ET, van Dongen JM, Juch JNS, Groeneweg JG, Kallewaard JW, de Boer MR, Koos B, Verhagen AP, Huygen FJP, van Tulder MW, Ostelo RWJG. Randomized controlled trials reflected clinical practice when comparing the course of low back pain symptoms in similar populations. J Clin Epidemiol 2019;116:122–32.
[14] McWilliams LA, Cox BJ, Enns MW. Mood and anxiety disorders associated with chronic pain: an examination in a nationally representative sample. PAIN 2003;106:127–33.
[15] Nair SC, Kievet W, Janse RW, Bijlsma JWJ, Fransen J, Lafeber FPJ, Welsing PMJ. Generalization and extrapolation of treatment effects from clinical studies in rheumatoid arthritis. Arthritis Care Res 2014;66:998–1007.
[16] Oliveira DS, Véila Ferreira Mendonça L, Sofia Monteiro Sampaio R, de Castro-Lopes J, Azevedo LFR. The impact of anxiety and depression on the outcomes of chronic low back pain multidisciplinary pain management—a multicenter prospective cohort study in pain clinics with one-year follow-up. Pain Med 2019;20:736–46.
[17] Page MJ, McKenzie JE, Bossuyt PM, Bouton I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akt E, Brennan SE, Chou R, Glanville J,
