Naltrexone 6 mg once daily versus placebo in women with fibromyalgia: a randomised, double-blind, placebo-controlled trial

Karin Due Bruun, Robin Christensen, Kirstine Amris, Henrik Bjarke Vaegter, Morten Rune Blichfeldt-Eckhardt, Lars Bye-Møller, Anders Holsgaard-Larsen, Palle Toft

Summary

Background Low-dose naltrexone is used to treat fibromyalgia despite minimal evidence for its efficacy. This trial aimed to investigate whether 12-week treatment with 6 mg low-dose naltrexone was superior to placebo for reducing pain in women with fibromyalgia.

Methods We did a single-centre, randomised, double-blind, placebo-controlled trial in Denmark. We enrolled women aged 18–64 years who were diagnosed with fibromyalgia. Participants were randomly assigned 1:1 to receive low-dose naltrexone (6 mg) or an identical-appearing placebo, using a computerised algorithm with no stratifications applied. Participants, investigators, outcome assessors, and statistical analysts were all masked to treatment allocation. The primary outcome was change in pain intensity on an 11-point numeric rating scale from baseline to week 12, in the intention-to-treat population. Safety was assessed in participants in the intention-to-treat population who received at least one dose of their allocated intervention. This trial was registered with ClincalTrials.gov (NCT04270877) and EudraCT (2019-000702-30).

Findings We screened 158 participants for eligibility from Jan 6, 2021, to Dec 27, 2022, and 99 patients were randomly assigned to low-dose naltrexone (n=49) or placebo (n=50). The mean age was 50·6 years (SD 8·8), one (1%) of 99 participants was Arctic Asian and 98 (99%) were White. No participants were lost to follow-up. The mean change in pain intensity was –1·3 points (95% CI –1·7 to –0·8) in the low-dose naltrexone group and –0·9 (–1·4 to –0·5) in the placebo group, corresponding to a between-group difference of –0·34 (–0·95 to 0·27; p=0·27, Cohen's d 0·23). Discontinuations due to adverse events were four (8%) of 49 in the low-dose naltrexone group and three (6%) of 50 in the placebo group. 41 (84%) of 49 patients in the low-dose naltrexone group had an adverse event versus 43 (86%) of 50 in the placebo group. One serious adverse event occurred in the placebo group and no deaths occurred.

Interpretation This study did not show that treatment with low-dose naltrexone was superior to placebo in relieving pain. Our results indicate that low-dose naltrexone might improve memory problems associated with fibromyalgia, and we suggest that future trials investigate this further.

Funding The Danish Rheumatism Association, Odense University Hospital, Danielsen's Foundation, and the Oak Foundation.

Copyright © 2023 Elsevier Ltd. All rights reserved.

Lancet Rheumatol 2024; 6: e31–39

Published Online December 5, 2023 https://doi.org/10.1016/S2665-9913(23)00278-3

See Comment page e5

Pain Research Group (K Due Bruun MD, Prof H B Vaegter PhD, M R Blichfeldt-Eckhardt PhD) **and Patient Panel** (L Bye-Møller), **Pain Center, Odense University Hospital, Odense, Denmark; Department of Orthopedics and Traumatology** (Prof A Holsgaard-Larsen PhD) **and Department of Anesthesiology and Intensive Care** (Prof P Toft DMSc), **Odense University Hospital, Odense, Denmark; Department of Clinical Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark** (K D Bruun, Prof R Christensen PhD, Prof H B Vaegter, M R Blichfeldt-Eckhardt, Prof A Holsgaard-Larsen, Prof P Toft); **The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark** (K D Bruun, Prof R Christensen, K Amris DMSc); **Department of Rheumatology, Bispebjerg and Frederiksberg Hospital, Frederiksberg, Denmark** (K Amris)

Correspondence to: Dr Karin Due Bruun, Pain Research Group, Pain Center, Odense University Hospital, Odense 5000, Denmark **karin.due.bruun@rsyd.dk**

Introduction

Fibromyalgia is a common debilitating condition affecting about 2% of the general population worldwide,[1] with a more than 9-fold greater prevalence among women in diagnosed populations.[2] Recent prevalence studies using new symptom-based diagnostic criteria show a more even ratio between sexes in general populations.[2] This discrepancy seems to reflect a severe under-diagnosis of fibromyalgia among men in patient populations.[3] Fibromyalgia syndrome is characterised by widespread pain and tenderness accompanied by a range of non-pain symptoms such as fatigue, sleep disturbance, and dyscognition. Pain in fibromyalgia has been shown to be related to alterations in functional connectivity in brain regions involved in pain processing, decreased activity in anti-nociceptive pathways, and increased activity in pro-nociceptive pathways.[4] The cause is still poorly understood but seems multifactorial, with different central and peripheral mechanisms as possible reinforcements of altered pain processing.[5]

There are several guideline-recommended pharmacological treatment options for fibromyalgia,[6] of which duloxetine, milnacipran, and pregabalin have been approved by the US Food and Drug Administration.[6] However, response rates to these treatments are low, dropouts are common because of side effects, and the European Medicines Agency has not approved these treatments because of the non-advantageous risk–benefit profile.[7–9] Non-pharmacological treatments, such as patient education, cognitive behavioural therapy, exercise, or multidisciplinary treatment, can improve pain and other fibromyalgia symptoms. The treatment effects seem to be

Research in context

Evidence before this study

We searched Medline for papers published in peer-reviewed journals from database inception to May 25, 2023, using the terms "naltrexone" and "fibromyalgia". We identified 46 articles, hereof two studies publishing results from clinical trials investigating the efficacy of low-dose naltrexone compared with placebo in patients with fibromyalgia. Both studies applied a cross-over design and used a dose of 4·5 mg. The first study was a single-blind pilot trial (ten women), and the second was a randomised placebo-controlled trial (31 women). These two trials indicated that low-dose naltrexone might be more effective than a placebo in reducing pain intensity in women with fibromyalgia. However, both studies were small and potentially biased due to several methodological weaknesses. In June 2023, a new trial with a cross-over design, testing a dose of 4·5 mg, and including 52 patients (46 women and 6 men) with fibromyalgia was published. This third study did not show an analgesic effect of low-dose naltrexone over a placebo. Several factors might have resulted in this negative result, for example the intended sample size of 140 participants was not reached and a sample size calculation was not provided for the pain outcome.

Added value of this study

The FINAL trial is the first randomised, double-blind, placebo-controlled trial with a parallel group design to investigate the efficacy of naltrexone 6 mg in women with fibromyalgia. Low-dose naltrexone was not superior to placebo in reducing pain at the group level. Among the secondary outcomes, we found a significant improvement only for memory problems related to fibromyalgia in favour of low-dose naltrexone. Discontinuations due to adverse events were low in both groups, and no concerns with safety related to treatment with this relatively high dose of 6 mg were seen.

Implications of the available evidence

According to the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials guidelines, many factors must be considered when evaluating the clinical importance of group differences, including responder analyses, secondary outcomes, and safety. A higher proportion of participants in the low-dose naltrexone group (45%) reported a more than 30% decrease in pain, than in the placebo group (28%). However, our study was not powered to detect a difference between groups regarding responder indices, and our sample size was most likely too small to detect a significant difference. Among the other key secondary outcomes, we found a significant between-group difference regarding the improvement of memory problems in favour of low-dose naltrexone treatment. The clinical relevance of this finding remains to be explored. We recommend more extensive trials with robust methods before definitive conclusions can be made about the clinical efficacy of low-dose naltrexone for treating fibromyalgia.

stable up to 14 weeks after the end of treatment, but then begin to decline.[10] Fibromyalgia is associated with a high symptom burden, increased use of health-care resources, work disability, and lower health-related quality of life than patients with other chronic diseases.[11] Thus, effective and safe treatment options are highly warranted.

Naltrexone is a non-selective opioid receptor antagonist that was marketed in the 1980s as an additional therapy for preventing relapse in patients with previous abuse of opioids or alcohol.[12] Low-dose naltrexone has been used as an off-label treatment for fibromyalgia for several years despite no evidence from large randomised controlled trials.[13] Low-dose naltrexone traditionally refers to doses of 1–5 mg,[14] however, in clinical practice, doses of up to 9 mg of naltrexone have been used to treat fibromyalgia.[15] Putative mechanisms of action of low-dose naltrexone could be a feedback-mediated increased expression of opioid receptors and opioid peptides with possible improvement of pain inhibition mediated via the endorphin system[16] or an anti-inflammatory effect mediated through antagonistic action at the Toll-like receptor 4 that is located on neuroimmune cells.[17]

Before initiating a clinical trial, we systematically searched the literature and found two small clinical trials conducted by the same research group investigating the efficacy of low-dose naltrexone compared with a placebo for treating pain in women with fibromyalgia.[18,19] Both studies used a dose of 4·5 mg and applied a cross-over design. The first trial was a single-blind pilot trial with ten participants.[18] The second trial was a double-blind, placebo-controlled, randomised trial that included 31 women with fibromyalgia.[19] The studies showed preliminary evidence that low-dose naltrexone might be superior to placebo in relieving pain and other symptoms of fibromyalgia. In the randomised control trial, no difference in overall tolerability was found, but headaches and vivid dreams were reported more frequently during treatment with low-dose naltrexone. Due to methodological weaknesses described in detail in a recent review,[13] both trials had a high risk of bias. Thus, we found the need for a larger and methodologically more robust randomised control trial to assess the potential efficacy of low-dose naltrexone for treating pain in patients with fibromyalgia. A new trial including 52 patients (46 women and 6 men) with fibromyalgia has recently been published.[20] As in the two earlier trials, a cross-over design and a 4·5 mg dose were used. This third study did not show an analgesic effect of low-dose naltrexone over a placebo.

The primary objective of the Fibromyalgia and Naltrexone (FINAL) study was to investigate whether 12-week treatment with naltrexone 6 mg was superior to

placebo in reducing the average pain intensity (during the past 7 days) in women with fibromyalgia. Secondary objectives included core fibromyalgia domains such as non-pain symptoms, daily functioning, health-related quality of life, global impression of change, and responder indices.

Methods

Study design

The FINAL study was a single-centre, randomised, double-blind, placebo-controlled superiority trial conducted at a tertiary pain rehabilitation centre in Denmark (Pain Center South, Odense University Hospital). The study was approved by the Ethical Committee of Southern Denmark (S-20190133) and the Danish Health and Medicines Authority (19/26406) and was reported to the Danish Data Protection Agency. The study was registered with the European Union Drug Regulation Authorities Clinical Trials Database (EudraCT-nr: 2019-000702-30), and the protocol was uploaded to ClinicalTrials.gov (NCT04270877) before the initiation of the study. A detailed protocol paper was published before the end of inclusion.[21] The original protocol is included in the appendix (pp 5–35). The justification for using a test dose of 6 mg was based on clinical practice and data from our previously published dose-response study,[22] in which we tested doses between 2·25 mg and 6 mg and found that doses higher than 4·5 mg, as used in previous trials, might be more efficacious without causing more harm.

Participants

Participants were recruited from the study site and through advertisements in national patient association magazines (both printed and internet-based). To be eligible, participants had to be women aged 18–64 years with fibromyalgia, and no history of neurological disease, inflammatory rheumatic disease, or active cancer. To confirm the fibromyalgia diagnosis, participants were required to fulfil the American College of Rheumatology 1990 criteria for fibromyalgia.[23] Pain had to be at least moderate in intensity, defined as an average pain score during the past week of at least four on a 0–10 numeric rating scale (NRS). Participants were allowed to continue their usual care and pain medication. Because of the interaction between opioids and naltrexone, participants were excluded if they had used opioids less than 4 weeks before entering the trial. Using opioids during the trial was considered a protocol violation. As one of the exploratory outcomes was an assessment of inflammation biomarkers, anti-inflammatory medication and non-steroidal anti-inflammatory drugs were not allowed 4 weeks before and during the trial. A complete list of eligibility criteria is available in the original protocol (appendix pp 22–23) and the published protocol.[21] Written informed consent was obtained from all subjects entering the study.

Randomisation and masking

Using a 1:1 allocation, participants were randomly assigned to treatment with 6 mg naltrexone or an identically-appearing placebo, using a computerised algorithm; no stratifications were applied. A data manager without involvement in the study made a sequential randomisation list based on permuted blocks of two to six individuals. The allocation was concealed in a password-protected computer file that was only accessible by the data manager. The primary investigator enrolled the participants and assigned them a sequential randomisation number, allocating them to one of the two groups. Participants, investigators, outcome assessors, and statistical analysts were all masked to the allocation and the permuted blocking strategy. A blinded interpretation was made before unmasking and is available in the appendix (pp 59–64).

Procedures

Tablets containing 1·5 mg naltrexone and identically appearing placebo tablets were manufactured at Glostrup Pharmacy (Glostrup, Denmark; an independent compounding pharmacy). The trial medication was shipped to Hospital Pharmacy Funen (Odense, Denmark), which received a copy of the randomisation list and blinded the medicine using identical cans labelled with the randomisation numbers. The timeframe for the study was 16 weeks, consisting of a 12-week treatment period (including a 4-week titration phase) and a 4-week washout period aiming to observe possible withdrawal symptoms (weeks 13 to 16). All participants started with one daily oral dosage of 1·5 mg low-dose naltrexone or placebo. During the 4-week titration phase, the dose was increased by one tablet per day each week to 4 tablets per day at week 4. Dose escalation was based on safety and tolerability, and delayed increments were allowed in case of unacceptable side effects. After the end of week 4, a maintenance dose was determined, equivalent to the highest dose tolerated at this timepoint. The trial medicine was taken once daily in the evening.

See Online for appendix

Due to issues related to the COVID-19 pandemic, the Danish Medicines Authority demanded that all Danish trials take appropriate actions to reduce the risk of infection. Therefore, it was decided to convert three follow-ups (ie, at weeks 4, 8, and 16) to telephone visits. The detailed visit schedule is available in the protocol (appendix p 25).

Outcomes

All patient-reported outcomes were measured at baseline and after 4, 8, and 12 weeks of treatment in a repeated measures design, with the 12-week assessment being of primary interest. The primary outcome measure was change in pain intensity from baseline to 12 weeks, using the level of pain question from the Fibromyalgia Impact Questionnaire-Revised (FIQR) questionnaire,[24] which measures the average pain within the past 7 days on an 11-point NRS, ranging from

0 indicating no pain to 10 indicating unbearable pain. This was measured in the intention-to-treat population. To reduce recall bias, all participants were asked to report their pain in the past 24 h in a handwritten diary, 7 days before baseline, and 7 days before week 8 and week 12.

Key secondary outcomes included the Patient's Global Impression of Change on a 1–7 verbal rating scale, the global impact of fibromyalgia by the FIQR total score, the Widespread Pain Index from the 2016 diagnostic criteria for fibromyalgia,[25] FIQR-tenderness item, FIQR-fatigue item, FIQR-sleep disturbance item, FIQR-depression item, FIQR-anxiety item, FIQR-memory problems item, FIQR-stiffness item, and FIQR physical function domain. The EQ-5D-5L[26] assessed health-related quality of life, including domains of mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. The EQ Visual Analogue Scale was used to determine the change in global health-related quality of life. The pressure pain threshold was measured three times at two points using a handheld pressure algometer (Somedic, Hörby, Sweden) and is reported as the average of the six measurements. The two points measured are right quadriceps 15 cm from apex patella and left trapezius 10 cm acromion (between acromion and C6 and C7). Each point is measured three times; point two and point three is measured 1 cm above and 1 cm below the first point. The pressure pain threshold was measured at baseline and after 12 weeks of treatment.

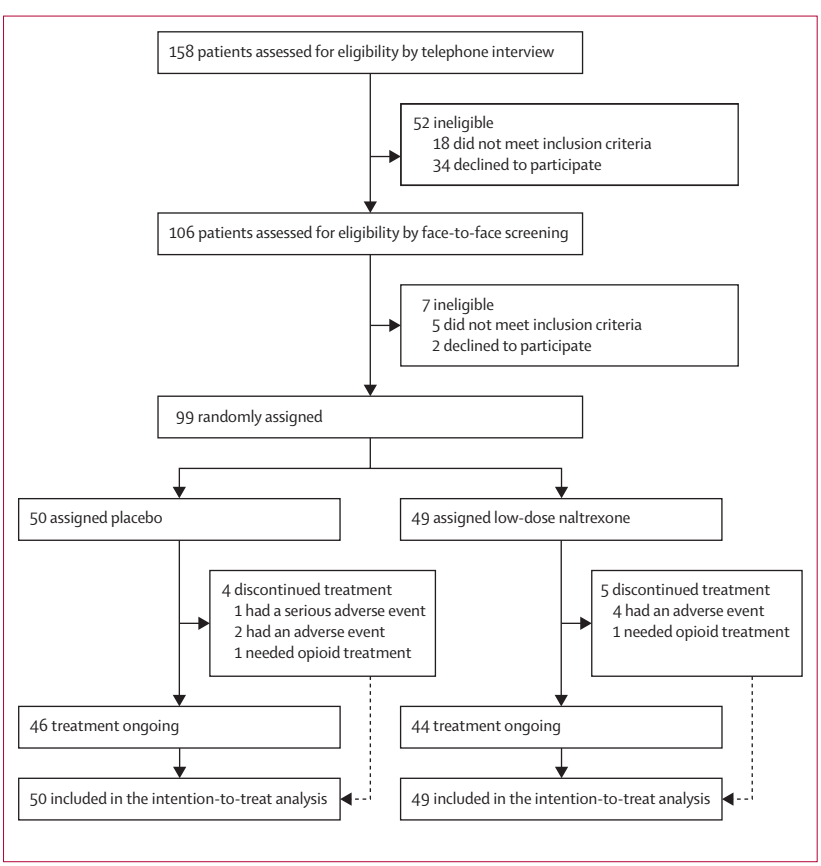

Figure 1: **Trial profile**

Other secondary supportive outcomes included investigating the number of responders in both treatment groups. Three responder categories were defined a priori as the number of responders with more than 15%, 30%, and 50% improvement in the primary outcome measure from baseline to 12 weeks.

For the surveillance of harms, both active and passive methods were used. The participants were encouraged to report adverse events spontaneously and were also asked about 12 common side effects via a questionnaire. If the participant reported harms categorised as grade 2 or higher, they were advised by the primary investigator to lower the dose. If harms were classified as grade 1, the decision about dosing was made individually in an agreement between the primary investigator and the participant. The primary investigator followed up with the participants by telephone until adverse events had ceased or were stable. The Common Terminology Criteria for Adverse Events (version 5.0) was used to grade the severity of harm.

Statistical analyses

Using data from our previous dose-response study, we estimated the self-reported pain intensity on a 0–10 NRS at baseline to have a mean of 6·7 points (SD 1·5) in the target population. According to the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) guidelines,[27] a minimal clinically important change is defined as a 15% decrease in pain (approximately 1·0 NRS point). In contrast, a 30% decrease (about 2·0 NRS points) is defined as a clinically meaningful change, and a 50% decrease is considered a substantial improvement. No definition of a minimal clinically important difference between groups is available from the IMMPACT guidelines. Using an estimated minimal clinically important difference between groups of 1·0, a SD of 1·5 (corresponding to a Cohen's effect size of 0·67), a statistical power of at least 80%, and a two-sided statistical significance level of 0·05, 74 patients were required for the intention-to-treat population (ie, 37 patients in each group). Expecting some attrition and drop-outs during the 12-week trial period, we decided to include 100 patients (ie, 50 patients in each group), potentially corresponding to a statistical power of more than 90% to detect a difference between groups in the intention-to-treat population. The statistical analysis plan was published online at ClinicalTrials.gov before the end of inclusion and is available in the appendix (pp 36–58).

Our main analyses comprised estimations of between-group differences in the continuous outcomes after 12 weeks for primary and secondary outcomes. Repeated measurements (T=0, 4, 8, and 12 weeks from baseline) were used in a linear mixed-effects model. The treatment

group, week, and the interaction between them were used as fixed effect factors, and participant identification as a random-effect parameter. All between-group differences based on the least square means were adjusted for baseline level to reduce the random variation. All p values and 95% CIs were two-sided. The main analyses were based on the intention-to-treat population, which included all participants assessed and randomly assigned at baseline. Using mixed effects models, missing data would be handled indirectly and statistically modelled using repeated-measures linear mixed models; mixed effects models are valid, assuming data are missing at random.[28]

We also calculated the number of responders (binary endpoints) in the two groups, based on participants who reported a more than 15%, 30%, and 50% decrease in pain after 12 weeks of treatment with low-dose naltrexone or placebo. These outcomes were analysed and reported as Risk Ratios (RR) with 95% CI comparing the proportions responding in the two groups.

To confirm the robustness of the main findings, sensitivity analyses were performed and reported on the main analyses for the per protocol population, ie, participants with at least 80% adherence to the prescribed treatment.

For ease of interpretation, and in line with EQ-5D reporting guidelines, the EQ-5D domains were dichotomised into the number and proportions of participants having no or slight problems (level 1–2) versus moderate, severe, or extreme problems (level 3–5). These dichotomous outcomes are reported for both groups at baseline and after 12 weeks of treatment, with no comparative statistics. The between-group differences for the continuous EQ Visual Analogue Scale outcome was assessed with comparative statistics as described previously.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Participants were recruited from Jan 6, 2021, to Dec 27, 2022. 158 patients were screened for eligibility; telephone interviews excluded 52, and another seven were excluded by face-to-face screening (figure 1). The remaining 99 eligible patients were randomly assigned to treatment with low-dose naltrexone (n=49) or placebo (n=50). One (1%) of 99 participants was Arctic Asian, and 98 (99%) were White, and the mean age was 50·6 years (SD 8·8; table 1).

Four (8%) of 49 participants in the low-dose naltrexone group and three (6%) of 50 in the placebo group discontinued treatment after week 4 because of intolerable side effects. One protocol violation occurred in each group due to non-related adverse events requiring opioid

	Low-dose naltrexone (n=49)	Placebo (n=50)	Total population (n=99)
Age, years	50·8 (8·8)	50·4 (8·9)	50·6 (8·8)
Ethnicity			
White	48 (98%)	50 (100%)	98 (99%)
Arctic Asian	1 (2%)	0	1 (1%)
Height, m	167·9 (7·1)	167·9 (5·8)	167·9 (6·4)
Bodyweight, kg	86·3 (17·1)	81·8 (15·5)	84·0 (16·4)
Body mass index, kg/m²	30·6 (5·8)	29·0 (5·1)	29·8 (5·5)
Duration of chronic pain, years	19·3 (12·3)	21·8 (11·3)	20·6 (11·8)
Pain medication			
None	5 (10%)	4 (8%)	9 (9%)
One	32 (65%)	33 (66%)	65 (66%)
Two or more	12 (24%)	13 (26%)	25 (25%)
Concomitant pain medication			
Paracetamol	42 (86%)	43 (86%)	86 (87%)
Tricyclic antidepressants or serotonin-noradrenalin-reuptake-inhibitor	12 (24%)	8 (16%)	20 (20%)
Gabapentin or pregabalin	3 (6%)	6 (12%)	9 (9%)
Other	2 (4%)	4 (8%)	6 (6%)
Level of pain,* average past 7 days (range 0–10)	6·3 (1·3)	6·2 (1·6)	6·3 (1·5)
Fibromyalgia impact questionnaire–revised total score (range 0–100)	55·2 (12·5)	54·1 (14·8)	54·6 (13·7)
Pain distribution, widespread pain index (range 0–19)	12·6 (3·3)	11·4 (4·0)	12·0 (3·7)
Level of tenderness,* average past 7 days (range 0–10)	6·1 (2·6)	5·9 (2·9)	6·0 (2·7)
Tenderness, average pressure pain threshold,† in kPa	188·4 (69·7)	198·6 (88·8)	193·5 (79·6)
Level of fatigue,* average past 7 days (0–10)	6·9 (1·5)	6·9 (1·7)	6·9 (1·6)
Level of sleep disturbance,* average past 7 days (range 0–10)	8·1 (1·7)	7·5 (2·0)	7·8 (1·9)
Level of depression,* average past 7 days (range 0–10)	2·8 (2·4)	2·9 (2·5)	2·9 (2·5)
Level of anxiety,* average past 7 days (range 0–10)	1·8 (2·7)	2·4 (2·7)	2·1 (2·7)
Level of memory problems,* average past 7 days (0–10)	6·2 (2·2)	5·2 (2·1)	5·7 (2·2)
Level of stiffness,* average past 7 days (range 0–10)	6·6 (1·9)	6·6 (2·1)	6·6 (2·0)
Fibromyalgia impact questionnaire–revised function domain (range 0–90)	47·4 (17·3)	50·2 (17·4)	48·8 (17·3)
EuroQoL 5 dimensions 5 levels‡			
Mobility; moderate, severe, or extreme	19 (39%)	23 (46%)	42 (42%)
Self-care; moderate, severe, or extreme	15 (31%)	10 (20%)	25 (25%)
Activity; moderate, severe, or extreme	33 (67%)	36 (72%)	69 (70%)
Pain; moderate, severe, or extreme	47 (96%)	44 (88%)	91 (92%)
Anxiety; moderate, severe, or extreme	8 (16%)	6 (12%)	14 (14%)
EuroQoL visual analog scale (range 0–100)	45·1 (17·5)	45·2 (17·3)	45·2 (17·3)

Data are n (%) or mean (SD) unless otherwise stated. *Fibromyalgia impact questionnaire–revised item. †Measured using a handheld pressure algometer. An average of the six measurements is reported. ‡The percentages represent participants reporting moderate or worse symptoms.

Table 1: **Baseline characteristics of the intention-to-treat population**

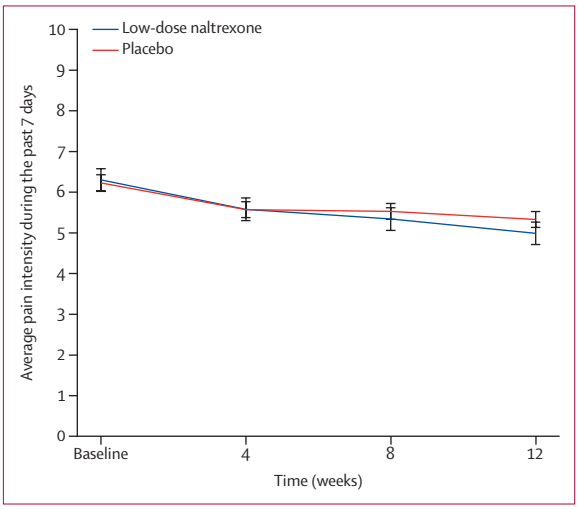

Figure 2: **Pain trajectory**
The trajectory for the average pain during the past 7 days (using the pain question from the Fibromyalgia Impact Questionnaire, revised) over time from baseline to the primary endpoint after 12 weeks.

treatment. No participants were lost to follow-up, and the primary outcome was assessed for the entire intention-to-treat population. For the per protocol population (n=90), a maximum maintenance dose of 6 mg was obtained in 35 (80%) of 44 participants in the low-dose naltrexone group versus 39 (85%) of 46 participants in the placebo group. A lower maintenance dose was obtained in nine participants in the low-dose naltrexone group (eight on 4·5 mg and one on 3 mg) and seven participants in the placebo group (three on 4·5 mg and four on 3 mg).

The within-group mean change in pain intensity (the primary outcome) was –1·3 NRS (95% CI –1·7 to –0·8) for the low-dose naltrexone group and –0·9 NRS (–1·4 to –0·5) for the placebo group. There was no significant difference between groups; the between-group difference was –0·34 NRS (95% CI –0·95 to 0·27; p=0·27), corresponding to a Cohen's *d* of 0·23. Based on the least square means (and standard errors [SEs]) the pain intensity measure trajectories are presented for both groups in figure 2. Table 2 lists the changes for the primary and secondary continuous outcomes for each group and the corresponding between-group differences after 12 weeks of treatment, with 95% CI and p values.

	Change from baseline to after 12 weeks of treatment		Between group differences (95% CI)	p value
	Low-dose naltrexone (n=49)	Placebo (n=50)		
Primary outcome				
Pain intensity,* NRS 0–10	–1·3 (–1·7 to –0·8)	–0·9 (–1·4 to –0·5)	–0·34 (–0·95 to 0·27)	0·27
Key secondary outcomes				
Global impression of change, median (IQR)	5 (4 to 6)	4 (4 to 5)	NA	0·20
Impact of fibromyalgia, Fibromyalgia Impact Questionnaire (revised) total score 0–100	–10·8 (–13·8 to –7·8)	–8·3 (–11·3 to –5·3)	–2·50 (–6·73 to 1·72)	0·24
Pain distribution, Widespread Pain Index 0–19	–2·4 (–3·3 to –1·4)	–1·7 (–2·6 to –0·8)	–0·64 (–1·95 to 0·67)	0·34
Level of tenderness,* NRS 0–10	–1·3 (–1·8 to –0·8)	–1·1 (–1·5 to –0·6)	–0·24 (–0·92 to 0·43)	0·48
Average pain pressure threshold†	2·6 (–12·5 to 17·7)	–9·1 (–23·9 to 5·7)	11·70 (–9·41 to 32·81)	0·28
Level of fatigue,* NRS 0–10	–1·0 (–1·4 to –0·5)	–0·9 (–1·4 to –0·5)	–0·04 (–0·69 to 0·60)	0·90
Level of sleep disturbance,* NRS 0–10	–1·7 (–2·3 to –1·2)	–1·6 (–2·2 to –1·0)	–0·16 (–0·99 to 0·68)	0·71
Level of depression,* NRS 0–10	–0·6 (–1·1 to –0·1)	–0·4 (–0·9 to 0·1)	–0·18 (–0·86 to 0·50)	0·61
Level of anxiety,* NRS 0–10	–0·3 (–0·6 to 0·1)	–0·4 (–0·8 to –0·1)	0·18 (–0·32 to 0·67)	0·49
Level of memory problems,* NRS 0–10	–1·4 (–1·9 to –1·0)	–0·5 (–0·9 to –0·1)	–0·93 (–1·57 to –0·30)	0·004
Level of stiffness,* NRS 0–10	–1·2 (–1·6 to –0·7)	–1·1 (–1·5 to –0·6)	–0·13 (–0·76 to 0·51)	0·70
Physical function, Fibromyalgia Impact Questionnaire (revised) function domain 0–90	–7·3 (–10·7 to –4·0)	–5·7 (–9·0 to –2·4)	–1·63 (–6·33 to 3·07)	0·50
Health-related quality of life, EuroQoL Visual Analog Scale 0–100	6·6 (2·2 to 11·0)	5·3 (0·9 to 9·7)	1·33 (–4·89 to 7·55)	0·68
Responder indices				
15% improvement in pain, n (%)	26 (53%)	21 (42%)	RR=1·26 (0·83 to 1·92)	0·27
30% improvement in pain, n (%)	20 (41%)	13 (26%)	RR=1·57 (0·88 to 2·79)	0·12
50% improvement in pain, n (%)	12 (24%)	7 (14%)	RR=1·75 (0·75 to 4·07)	0·19

Repeated measures mixed effects models: estimates are presented as least squares means (95% CI) per group, and the difference between them is reported with the corresponding 95% CI–unless otherwise stated. NRS=Numeric Rating Scale. RR=relative risk. *Items from the Fibromyalgia Impact Questionnaire (revised). †Measured using a handheld pressure algometer. An average of the six measurements is reported in KPa.

Table 2: **Primary, key secondary, and other secondary outcomes at 12 weeks from baseline based on the intention-to-treat population**

There was no significant difference between groups for most of the secondary outcomes. Across all the secondary continuous outcomes, we only found a significant difference between the groups for memory problems in favour of low-dose naltrexone (–0·93, 95% CI –1·57 to –0·30, p=0·004). When adjusting for multiplicity (0·05÷16=0·003), this difference lost its significance. The Patient's Global Impression of Change for both groups shows that more participants in the low-dose naltrexone group reported an overall improvement than the placebo group (appendix p 2). However, a statistically significant difference in Patient's Global Impression of Change score was not observed (table 2).

A 15% reduction in pain was seen in 26 (53%) of 49 women in the low-dose naltrexone group and 21 (42%) of 50 women in the placebo group, corresponding to a relative risk (RR) of responding of 1·26 (95% CI 0·83 to 1·92, p=0·27; table 2). The number of participants who reported a clinically meaningful change (at least 30% pain reduction) was 20 (41%) in the low-dose naltrexone group and 13 (26%) in the placebo group (RR 1·57 [95% CI 0·88 to 2·79], p=0·12). For participants with at least 50% pain reduction (defined as a substantial change), the numbers were 12 (24%) in the low-dose naltrexone group and seven (14%) in the placebo group (RR 1·75 [95% CI 0·75 to 4·07], p=0·19).

Sensitivity analyses showing the between-group differences for the per protocol population are available in the appendix (p 4). For the primary outcome (change in pain intensity), the between-group difference was larger in the per protocol population (–0·47 NRS, 95% CI –1·11 to 0·18; p=0·15) compared with the intention-to-treat population. Regarding the number of pain responders, the RR was slightly larger in the per protocol population (eg, for 30% pain responders, RR 1·61 [95% CI 0·92 to 2·82], p=0·09). The change in memory problems remained statistically significant (–1·01, 95% CI –1·69 to –0·34; p=0·004).

The dichotomised EQ-5D domains are available in the appendix (p 3). For the pain and discomfort domain, most participants in both groups reported problems as level 3–5. A change in category from level 3–5 to level 1–2 was observed in 12 (24%) of 49 in the low-dose naltrexone group versus 4 (8%) of 50 in the placebo group.

Adverse events in both groups are summarised in table 3, with a breakdown by grade of the event and the frequencies of 12 predefined adverse events. Adverse events were reported by 41 (84%) of 49 patients in the low-dose naltrexone group (19 [39%] of a moderate grade) and 43 (86%) of 50 in the placebo group (17 [34%] moderate). The median number of adverse events reported per patient was three in the low-dose naltrexone group and two in the placebo group. The most frequent adverse event was headache, which occurred in 18 (37%) patients in the low-dose naltrexone group and 19 (38%) in the placebo group. Vivid dreams, diarrhoea, constipation, increased appetite, dizziness, and hot flushes were reported more than twice as frequently in the low-dose naltrexone group than the placebo group. However, constipation and increased appetite were not commonly reported (<10%). One serious adverse event occurred in the placebo group (hospitalisation for 5 h due to severe abdominal pain). None of the reported adverse events were unexpected. No deaths occurred.

	Low-dose naltrexone (n=49)	Placebo (n=50)
Final dose, mg	6·0 (4·5–6·0)	6·0 (6·0–6·0)
Exposure time, patient weeks	12 (12–12)	12 (12–12)
Adverse events	41 (84%)	43 (86%)
Adverse events, n events (rate per patient)	3 (1–6)	2 (1–4)
Mild adverse events	39 (80%)	42 (84%)
Moderate adverse events	19 (39%)	17 (34%)
Serious adverse events	0	1 (2%)
Deaths	0	0
Pre-specified adverse events		
Headache	18 (37%)	19 (38%)
Vivid dreams	19 (39%)	9 (18%)
Diarrhoea	14 (29%)	7 (14%)
Constipation	8 (16%)	2 (4%)
Abdominal ache	11 (22%)	10 (20%)
Nausea	13 (27%)	14 (28%)
Increased appetite	5 (10%)	2 (4%)
Dizziness	14 (29%)	7 (14%)
Palpitations	2 (4%)	0
Hot flushes	16 (33%)	7 (14%)
Dry mouth	10 (20%)	10 (20%)
Depressed mood	2 (4%)	1 (2%)

Data are n (%) or median (IQR). The safety population was defined as participants in the intention-to-treat population who received at least one dose of their allocated intervention. The severity of an adverse event refers to the maximum intensity of the event. An event is considered mild if it does not interfere with activities of daily life, moderate if it limits instrumental activities of daily life, and severe if it interferes substantially with the patient's activities of daily life. An adverse event is classified as serious if fatal or life-threatening, requires inpatient hospitalisation, causes substantial disabling, or requires medical intervention to prevent permanent impairment or damage.

Table 3: **Adverse events**

Discussion

To the best of our knowledge, this trial is the first rigorously designed, conducted, and reported randomised study evaluating the efficacy of low-dose naltrexone 6 mg for 12 weeks compared with placebo for treating pain in women with fibromyalgia. We found that treatment with low-dose naltrexone was not superior to placebo for reducing the average pain intensity. Among the key secondary outcomes, we only found a significant between-group difference in improving memory problems; however, this finding might be a false positive due to multiplicity. The study revealed no concerns with harms related to treatment with this relatively high dose of 6 mg low-dose naltrexone.

A recent systematic review investigating the efficacy of low-dose naltrexone for treating fibromyalgia showed that two early placebo-controlled studies lacked scientific robustness, and their preliminary evidence of a positive effect was considered potentially biased.[13] Data have recently been published from a third trial conducted in Denmark, where several methodological issues were improved, eg, a priori sample size calculation, similar lengths of treatment periods (21 days), and inclusion of a wash-out period (14 days) between the placebo and the low-dose naltrexone conditions. In this third trial, the primary outcomes were mean changes in FIQR total score and Summed Pain Intensity Rating on a 0–30 NRS (summing three subscores of pain during rest 0–10, personal hygiene measures 0–10, and activities of daily living 0–10), measured as the average pain intensity during the past 3 days. The study did not show significant between-group differences for these two primary outcomes. The sample size calculation was based on Fibromyalgia Impact Questionnaire data from the early trials, and an estimate of a sample size adequate to detect a minimal clinically important difference in Summed Pain Intensity Rating was not provided, but the observed very small effect size (Cohen's *d* 0·04) indicated there was no clinically relevant difference for the change in Summed Pain Intensity Rating.

Our research group previously conducted a dose-response study, testing doses between 0·75 mg and 6 mg, providing an estimate of the effective dose in 50% of 3·88 mg and 95% of 5·4 mg.[22] However, a larger dose range might have given a higher estimate. As clinical practice has changed during the past decade with the use of doses of up to 9 mg of naltrexone,[15] combined with no safety concerns related to treatment with doses up to 6 mg in our dose-response trial, we decided to use a test dose of 6 mg. Acknowledging that one size does not fit all, we also chose to include a titration phase, allowing for delayed increments.

In the current trial, the observed effect size for pain improvement was small, and not significant (Cohen's *d* 0·23).[9] According to IMMPACT guidelines, there is a risk that clinically meaningful improvements for individual patients can be obscured by small mean group differences.[29] Thus, a benefit–risk evaluation at the study level is recommended, including evaluation of secondary outcomes, responder analysis, safety parameters, and a comparison with other available therapies.[29,30]

Across the key secondary outcomes, we found small improvements of all patient reported outcomes in both groups, with no significant between-group differences except for FIQR-memory problems. Whether this finding is a false positive due to multiplicity, remains to be explored. None of the previous low-dose naltrexone trials have included measures of memory problems or other measures of disturbed cognition as an outcome.[9,18–20]

In our sensitivity analysis, we found the number of 30% pain responders to be 20 (45%) of 44 in the low-dose naltrexone group and 13 (28%) of 46 in the placebo group, corresponding to a number needed to treat of 6. Our study was not powered to detect a significant difference regarding responder indices. However, when looking at the 95% CI around the estimand of 30% response rates, we hypothesise that this finding could be interpreted as a potential difference to be explored in future trials. Subgroups of patients with fibromyalgia might respond differently to low-dose naltrexone treatment, and we intend to conduct a responder analysis based on levels of inflammatory biomarkers and specific biomarkers of glial activation, hypothesising that an inflammatory subgroup might benefit from the treatment. Results will be published in subsequent papers.

Discontinuations due to adverse events were very low in our trial, with 4 (8%) of 49 in the low-dose naltrexone group and 2 (4%) of 50 in the placebo group. In two earlier low-dose naltrexone cross-over trials, the drop-out rates were about 10%. As a comparison, in a systematic review of serotonin and norepinephrine reuptake inhibitors the number of withdrawals due to adverse events was reported to be 19% in the serotonin and norepinephrine reuptake inhibitor group and 10% in the placebo group.[9]

The main strength of this study is the use of a robust method, aiming to reduce the level of bias. We used a computerised random sequence generation that kept participants, outcome assessors, investigators, and statisticians masked to the allocation. Although our study was a single-centre study, participants were recruited from all over Denmark. Recruitment through advertisements provided a representative sample of patients with varying impacts of the disease. The treatment groups were comparable regarding the baseline characteristics, and treatment compliance was high in both groups. No participants were lost to follow-up. The main analysis was based on the intention-to-treat population, and sensitivity analysis of the per protocol population showed similar results.

The study's main limitation is that it was only powered to detect a difference between groups of 1·0 NRS points for the intensity of pain. The inclusion of 13 secondary outcomes and three dichotomous responder index outcomes might have increased the chance of a positive finding among the secondary outcomes. Another limitation could be around the external validity of the trial. As we primarily included White women aged 18–64 years, our results cannot be generalised to men, adolescents, older adults, or other ethnic groups. Using a 12-week follow-up period, our study does not provide knowledge about long-term treatment or adverse effects.

In conclusion, the current study did not show that treatment with low-dose naltrexone was superior to placebo in reducing pain in women with fibromyalgia in general. Our results indicate that low-dose naltrexone might improve memory problems associated with fibromyalgia, and we suggest that future trials investigate this further.

Contributors
KDB was the primary investigator and participated in the conceptualisation, methodology, literature search, data visualisation, data interpretation, and the first draft of the manuscript. RC was the senior biostatistician who helped write the statistical analysis plan and did the statistical analyses, data visualisation, and interpretation. RC also participated in the protocol writing (conceptualisation and methodology), manuscript review, and editing. KA, HBV, MRB-E, and AH-L participated in the protocol writing (conceptualisation and methodology), statistical analysis plan, data interpretation, manuscript review, and editing. LB-M was the patient representative and participated in data interpretation, manuscript review, and editing. PT was the sponsor and participated in protocol writing (conceptualisation and methodology), statistical analysis plan, data interpretation, manuscript review, and editing. All authors had full access to all study data and were together responsible for the decision to submit the manuscript. KDB and RC accessed and verified the data.

Declaration of interest
We declare no competing interests.

Data sharing
The study protocol, the statistical analysis plan, and a blinded summary is available in the appendix (pp 5–64). De-identified participant data can be retrieved with the support of the primary investigator, preceded by a signed data access agreement form.

Acknowledgments
Study data were collected and managed using REDCap electronic data capture tools hosted at OPEN, Open Patient Data Explorative Network, Odense University Hospital, Region of Southern Denmark. Thanks to Claire Gudex, Department of Clinical Research, University of Southern Denmark, for commenting on and editing the manuscript. Operating expenses were covered by funding from the Danish Rheumatism Association (A6158), Odense University Hospital's PhD fund for operating expenses (A3650), and Aase og Ejnar Danielsen's Foundation. The Section for Biostatistics and Evidence-Based Research at the Parker Institute, Bispebjerg and Frederiksberg Hospital is supported by a core grant from the Oak Foundation (OCAY-18-774-OFIL).

References

1 Heidari F, Afshari M, Moosazadeh M. Prevalence of fibromyalgia in general population and patients, a systematic review and meta-analysis. *Rheumatol Int* 2017; **37:** 1527–39.
2 Wolfe F, Walitt B, Perrot S, Rasker JJ, Häuser W. Fibromyalgia diagnosis and biased assessment: sex, prevalence, and bias. *PLoS One* 2018; **13:** e0203755.
3 Bruun KD, Jensen HI, Blichfeldt-Eckhardt MR, et al. Performance of the 2016 diagnostic criteria for fibromyalgia in a tertiary care pain rehabilitation setting: a diagnostic accuracy study. *Scand J Pain* 2021; **22:** 67–76.
4 Cagnie B, Coppieters I, Denecker S, Six J, Danneels L, Meeus M. Central sensitization in fibromyalgia? A systematic review on structural and functional brain MRI. *Semin Arthritis Rheum* 2014; **44:** 68–75.
5 Pinto AM, Luís M, Geenen R, et al. Neurophysiological and psychosocial mechanisms of fibromyalgia: a comprehensive review and call for an integrative model. *Neurosci Biobehav Rev* 2023; **151:** 105235.
6 Tzadok R, Ablin JN. Current and emerging pharmacotherapy for fibromyalgia. *Pain Res Manag* 2020; **2020:** 6541798.
7 Lian YN, Wang Y, Zhang Y, Yang CX. Duloxetine for pain in fibromyalgia in adults: a systematic review and a meta-analysis. *Int J Neurosci* 2020; **130:** 71–82.
8 Tzellos TG, Toulis KA, Goulis DG, et al. Gabapentin and pregabalin in the treatment of fibromyalgia: a systematic review and a meta-analysis. *J Clin Pharm Ther* 2010; **35:** 639–56.
9 Welsch P, Üçeyler N, Klose P, Walitt B, Häuser W. Serotonin and noradrenaline reuptake inhibitors (SNRIs) for fibromyalgia. *Cochrane Database Syst Rev* 2018; **2:** CD010292.
10 Kundakci B, Kaur J, Goh SL, et al. Efficacy of nonpharmacological interventions for individual features of fibromyalgia: a systematic review and meta-analysis of randomised controlled trials. *Pain* 2022; **163:** 1432–45.
11 Lee JW, Lee KE, Park DJ, et al. Determinants of quality of life in patients with fibromyalgia: a structural equation modeling approach. *PLoS One* 2017; **12:** e0171186.
12 Sudakin D. Naltrexone: not just for opioids anymore. *J Med Toxicol* 2016; **12:** 71–75.
13 Partridge S, Quadt L, Bolton M, et al. A systematic literature review on the clinical efficacy of low dose naltrexone and its effect on putative pathophysiological mechanisms among patients diagnosed with fibromyalgia. *Heliyon* 2023; **9:** e15638.
14 Toljan K, Vrooman B. Low-dose naltrexone (LDN)-review of therapeutic utilization. *Med Sci (Basel)* 2018; **6:** 82.
15 Driver CN, D'Souza RS. Efficacy of low-dose naltrexone and predictors of treatment success or discontinuation in fibromyalgia and other chronic pain conditions: a fourteen-year, enterprise-wide retrospective analysis. *Biomedicines* 2023; **11:** 1087.
16 Metz MJ, Daimon CM, Hentges ST. Reported benefits of low-dose naltrexone appear to be independent of the endogenous opioid system involving proopiomelanocortin neurons and β-endorphin. *eNeuro* 2021; **8:** ENEURO.0087-21.2021.
17 Kučić N, Rački V, Šverko R, Vidović T, Grahovac I, Mršić-Pelčić J. Immunometabolic modulatory role of naltrexone in BV-2 microglia cells. *Int J Mol Sci* 2021; **22:** 8429.
18 Younger J, Mackey S. Fibromyalgia symptoms are reduced by low-dose naltrexone: a pilot study. *Pain Med* 2009; **10:** 663–72.
19 Younger J, Noor N, McCue R, Mackey S. Low-dose naltrexone for the treatment of fibromyalgia: findings of a small, randomized, double-blind, placebo-controlled, counterbalanced, crossover trial assessing daily pain levels. *Arthritis Rheum* 2013; **65:** 529–38.
20 Bested KJL, Andresen T, Tarp G, et al. Low-dose naltrexone for treatment of pain in patients with fibromyalgia: a randomized, double-blind, placebo-controlled, crossover study. *Pain Rep* 2023; **8:** e1088.
21 Bruun KD, Amris K, Vaegter HB, et al. Low-dose naltrexone for the treatment of fibromyalgia: protocol for a double-blind, randomized, placebo-controlled trial. *Trials* 2021; **22:** 804.
22 Bruun-Plesner K, Blichfeldt-Eckhardt MR, Vaegter HB, Lauridsen JT, Amris K, Toft P. Low-dose naltrexone for the treatment of fibromyalgia: investigation of dose-response relationships. *Pain Med* 2020; **21:** 2253–61.
23 Wolfe F, Smythe HA, Yunus MB, et al. The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multicenter Criteria Committee. *Arthritis Rheum* 1990; **33:** 160–72.
24 Bennett RM, Friend R, Jones KD, Ward R, Han BK, Ross RL. The Revised Fibromyalgia Impact Questionnaire (FIQR): validation and psychometric properties. *Arthritis Res Ther* 2009; **11:** R120.
25 Wolfe F, Clauw DJ, Fitzcharles MA, et al. 2016 revisions to the 2010/2011 fibromyalgia diagnostic criteria. *Semin Arthritis Rheum* 2016; **46:** 319–29.
26 Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: past, present, and future. *Appl Health Econ Health Policy* 2017; **15:** 127–37.
27 Dworkin RH, Turk DC, McDermott MP, et al. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. *Pain* 2009; **146:** 238–44.
28 Detry MA, Ma Y. Analyzing repeated measurements using mixed models. *JAMA* 2016; **315:** 407–08.
29 Dworkin RH, Turk DC, McDermott MP, et al. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. *Pain* 2009; **146:** 238–44.
30 Kleykamp BA, Dworkin RH, Turk DC, et al. Benefit–risk assessment and reporting in clinical trials of chronic pain treatments: IMMPACT recommendations. *Pain* 2022; **163:** 1006–18.