Fibromyalgia and Low Dose Naltrexone

Studies of Diagnostic Accuracy and Treatment Efficacy in a Specialised Pain Care Setting

Bruun, Karin Due

DOI:
10.21996/td0f-z434

Publication date:
2024

Document version:
Final published version

Document license:
Unspecified

Citation for pulished version (APA):
Bruun, K. D. (2024). *Fibromyalgia and Low Dose Naltrexone: Studies of Diagnostic Accuracy and Treatment Efficacy in a Specialised Pain Care Setting*. [Ph.D. thesis, SDU]. Syddansk Universitet. Det Sundhedsvidenskabelige Fakultet. https://doi.org/10.21996/td0f-z434

PhD Dissertation

Fibromyalgia and Low Dose Naltrexone

Studies of Diagnostic Accuracy and Treatment Efficacy in a Specialised Pain Care Setting

Karin Due Bruun, MD

Institute of Clinical Research
Faculty of Health Sciences
University of Southern Denmark

Author
Karin Due Bruun (former name: Karin Bruun Plesner)
Pain Center, Odense University Hospital
University of Southern Denmark

Main supervisor
Professor Palle Toft, MD, PhD, DMSc.
Department of Anesthesiology and Intensive Care, Odense University Hospital
University of Southern Denmark

Co-supervisors
Associate professor Kirstine Amris, MD, DMSc.
The Parker Institute, Bispebjerg and Frederiksberg Hospital
Copenhagen University

Professor Henrik Bjarke Vaegter, MSc, PhD.
Pain Center, Odense University Hospital
University of Southern Denmark

Associate professor Morten Rune Blichfeldt-Eckhardt, MD, PhD.
Pain Center, Middelfart Hospital
University of Southern Denmark

Assessment committee
Professor Hanne Lindegaard, MD, PhD (chair)
Department of Rheumatology, Odense University Hospital and University of Southern Denmark

Professor Ole Mathiesen, MD, PhD, DMSc.
Department of Anesthesiology, Copenhagen University and Zealand University Hospital

Professor Emmanuel Bäckryd, MD, PhD
Pain and Rehabilitation Center, Linköping University, Sweden

Preface

This PhD thesis is based on research performed during my employment at the Pain Center, Department of Anaesthesiology and Intensive Care, Odense University Hospital, and as a part-time PhD student at the Faculty of Health Sciences, University of Southern Denmark (2019-2023).
The thesis is based on the following papers:

Paper I
Performance of the 2016 diagnostic criteria for fibromyalgia in a tertiary pain rehabilitation setting: A diagnostic accuracy study.
Bruun KD, Jensen HI, Blichfeldt-Eckhardt MR, Vaegter HB, Toft P, Amris K, Kvorning N. Scand J Pain. 2021 Oct 20;22(1):67-76.

Paper II
Low Dose Naltrexone for the treatment of fibromyalgia: Investigation of dose-response relationships.
Bruun-Plesner K, Vaegter HB, Blichfeldt-Eckhardt MR, Amris K, Toft P.Pain Med. 2020 Oct 1;21(10):2253-2261.

Paper III
Low dose naltrexone for the treatment of fibromyalgia: Protocol for a double-blind, randomized, placebo-controlled trial.
Bruun KD, Amris K, Vaegter HB, Blichfeldt-Eckhardt MR, Holsgaard-Larsen A, Christensen R, Toft P. Trials. 2021 Nov 15;22(1):804.

Paper IV
Naltrexone 6 mg oral once daily versus placebo in women with fibromyalgia: a randomised, double-blind, placebo-controlled trial.
Bruun KD, Christensen R, Amris K, Vaegter HB, Blichfeldt-Eckhardt MR, Holsgaard-Larsen A, Bye-Moeller L, Toft P. Lancet Rheum. 2023 Dec 5; e-pub ahead of print.

Acknowledgements

Many people have been involved in my PhD process to whom I wish to express my gratitude. Firstly, I owe warm thanks to Gitte Handberg (former Head of the Pain Centre), who convinced me to apply for a position at the Pain Centre in Odense and gave me time for research even though I was not a researcher. Secondly, I wish to thank my principal supervisor, Professor Palle Toft, for supporting my idea of investigating the efficacy of low-dose naltrexone and talking me into a PhD thesis on the subject. I feel privileged to be allowed to do a PhD thesis relatively late in my professional career. For that, I wish to thank the head of my department for their indispensable support. When I started as a PhD student, I felt pretty competent as a physician but was a rookie in research. I had the privilege of picking out my team of co-supervisors, and I wish to thank you all for taking on the task and for patiently and competently teaching me the skills of research. You have all been a great support and inspiration during the past five years. Apart from being my supervisors, you have also been good colleagues. I specifically want to thank Kirstine Amris for being my dedicated mentor in fibromyalgia research, Morten Rune Blichfeldt-Eckhardt for our many exciting and nerdy talks about pain science, and Henrik Bjarke Vaegter for your everlasting moral support and your highly needed help on putting together the pieces in the dissertation.

Apart from my supervisors, I have received great help from several other researchers. Thanks to Robin Christensen and Anders Holsgaard Larsen for teaching me how to rigorously design a randomised controlled trial. We would not have made it to The Lancet Rheumatology without you! Thanks to Nina Kvorning for kickstarting the diagnostic accuracy study and all your dedicated hours of work on the study in Middelfart. Also, thanks to Hanne Irene Jensen for your valuable help on methodology in Study I and Jørgen Lauridsen for your indispensable help with the statistics in Study II. Many collaborative partners have contributed with support on the two drug trials, and grateful greetings go to Annie Sten and Henriette Kunoy from the GCP unit in Odense, Lars Søgaard from OPEN, Christian Østergaard Nielsen from Glostrup Pharmacy, Lene Sehested from Hospital Pharmacy Funen, and Marianne Augustenborg and Maria Dauerhøi from R-lab.

A big hug to Eva McGhee, a unique study nurse who lent me her expertise in GCP trials and created a warm and humorous patient environment. I have also received great help from the secretary at the Pain Centre. Vibeke Skovdal Albrechtsen and Benedikte Mølgaard deserve a big hug for helping me with booking and other practical issues. Warm greetings to all my dear colleagues at the Pain Centre for cheering on me and carrying over with me in periods of mental absence.

Special thanks to all the patients who participated in the trials. I am incredibly impressed by the dedication I met from all of you to contribute to research. Another very dedicated person is Lars Bye-Møller. Thank you, Lars, for your valuable assistance as our patient representative in study III. I also want to thank the Danish Rheumatism Association and the Danish Fibromyalgia and Pain Association for writing about my research and helping recruit participants.

I have appreciated many walks and talks about research and writing in the past five years with Nana Hyldig and Marie Mydtskov. Thanks to both of you for your sound advice and for listening to all

my complaints. To my sister-in-law, Dixi Louise Strand, huge thanks for the valuable assistance in proofreading essential parts of the dissertation.

Finally, and most importantly, I want to thank my husband, Jan Due. Thank you for always being available and engaged when I need you. Thank you for celebrating my victories and for your warm support during setbacks. You are the love of my life.

Financial support

I thank the below-mentioned organisations and foundations for their generous financial support.

My PhD salary was funded by:

- The Graduate School of Health Sciences, University of Southern Denmark
- The Region of Southern Denmark
- The Department of Anesthesiology and Intensive Care, Odense University Hospital

Operating expenses for Study II were funded by:

- The Danish Rheumatism Association

Operating expenses for Study III were funded by:

- The Danish Rheumatism Association
- Odense University Hospital's PhD foundation for operation expenses
- Aase and Ejnar Danielsen Foundation

Summary

Fibromyalgia is a common disorder associated with a high symptom burden and a low health-related quality of life, impacting daily functioning, working capacity, and social participation. Fibromyalgia is characterised by diffuse pain and tenderness related to healthy tissues and constitutes a prototype of nociplastic pain, where the pain is caused by augmented central pain processing. Its prevalence in the general population is about 2%, with a small female predominance. However, due to controversies and different beliefs among physicians, fibromyalgia is underdiagnosed in clinical patient populations, and the gender distribution is skewed, with more than 90% of the diagnosed being women. Diagnosing fibromyalgia timely and when appropriate can prevent unnecessary investigations and worries related to diagnosis uncertainty, support better self-care and guide a targeted pharmacological treatment strategy. In Denmark, patients with fibromyalgia can be referred to a specialised pain care centre where treatment is offered individually, combining pharmacological treatments with cognitive and behavioural therapeutic interventions. Guideline-recommended pharmacological treatments for fibromyalgia are centrally-acting drugs which influence neurotransmitters involved in central pain processing, e.g. pregabalin, which inhibits facilitatory neurotransmitters, or duloxetine, which promotes the release of inhibitory neurotransmitters. However, effect sizes are small, side effects are common, and only a minority of fibromyalgia patients benefit substantially from these treatments. During the last decade, an increased use of low-dose naltrexone (LDN) as an off-label treatment for fibromyalgia has been observed. A few small preliminary studies on the efficacy of LDN on fibromyalgia have shown promising results. However, these studies are potentially biased by several methodological weaknesses, with a risk of overestimating the effect. The aim of this thesis is twofold. Firstly, investigating how new survey-based diagnostic criteria for identifying fibromyalgia perform among patients with mixed chronic pain syndromes, and secondly, investigating the efficacy of LDN for treating fibromyalgia pain using robust methodology. The studies presented here include one diagnostic accuracy study (**Study I**) and two drug trials with LDN (**Study II** and **Study III**).

Study I (published in Scand J Pain 2021) investigated the diagnostic accuracy of the survey-based 2016 diagnostic criteria for fibromyalgia in a population of patients with mixed chronic pain syndromes referred to specialised pain care. No similar studies have previously been performed. The 2016 criteria showed a high sensitivity and an acceptable specificity in the present population, thus bringing new evidence that these criteria are valuable for the clinical identification of fibromyalgia and for identifying fibromyalgia for research purposes in specialised pain care settings.

Due to a skewed gender distribution among patients diagnosed with fibromyalgia, only women with fibromyalgia were included in the LDN trials. **Study II** (published in Pain Med 2020) investigated dose-response relationships using the “up-and-down” method, and results and experiences from **Study II** served to qualify the feasibility of **Study III**. **Study III** (published in Lancet Rheum 2023) investigated the clinical efficacy of 6 mg naltrexone on pain in women with fibromyalgia using a randomised, placebo-controlled, double-blind (RCT) design. Results from the RCT showed no general pain-relieving effect of LDN compared to placebo in women with fibromyalgia. Differences

regarding 30% pain responders approached statistical significance, indicating that there might be more pain responders to LDN than to placebo. Among the secondary outcomes, a significant positive effect on memory problems was observed in favour of LDN. No concerns were raised about safety. **Study III** was performed according to the highest standards for conducting and reporting clinical trials, represents the largest trial of its kind to date, and thus contributes substantially to the current knowledge about the efficacy of LDN for treating fibromyalgia.

To summarise, the findings from the current thesis support the use of the survey-based 2016 diagnostic criteria for fibromyalgia as an easy-to-apply tool for fibromyalgia identification in clinical practice and for research purposes in specialised pain care settings. The RCT could not demonstrate that LDN has a general effect on fibromyalgia pain. However, a trend towards more pain responders was found for the LDN group compared to placebo, and subgroups that benefit from LDN treatment might exist and call for further studies.

Dansk resume

Fibromyalgi er en lidelse hvor kerne symptomet er diffust udbredte smerter eller ømhed, som ikke kan forklares af sygdom i det smertefulde væv. Smerterne menes at være forårsaget af en ændret forarbejdning af smertesignaler i centralnervesystemet karakteriseret ved øget aktivitet i smerteførende nervefibre og i hjerneområder involveret i smertebearbejdning. Tilstanden påvirker ofte både funktionsniveau, arbejdsevne og social deltagelse og kan påvirke livskvaliteten væsentligt. Fibromyalgi er en almindelig tilstand, der findes hos ca. 2% af befolkningen, heraf er ca. 60% kvinder. På grund af forskellige holdninger til sygdommen blandt læger bliver diagnosen dog ofte ikke stillet. Dette gælder især mænd, og 90% af de diagnosticerede er derfor kvinder. Der er flere gode grunde til at diagnosticere fibromyalgi, når det er relevant. Dels kan en korrekt diagnose forebygge bekymringer hos patient og læge og hermed forhindre unødvendige undersøgelser. Desuden vil en diagnose kunne støtte bedre egenomsorg og guide den rette behandling. I Danmark kan patienter med fibromyalgi blive henvist til et Smertecenter, hvor behandlingen tilrettes individuelt, og hvor medicinske og psykologiske behandlingsformer kombineres. De medicinske behandlinger, der i dag anbefales til behandling af fibromyalgi, er præparater, der virker i centralnervesystemet, hvor de enten fremmer frigivelsen af smertedæmpende signalstoffer eller blokerer frigivelsen af smerteaktiverende signalstoffer. Desværre har disse præparater ikke en generelt god effekt på fibromyalgi. De kan være forbundet med mange bivirkninger, og kun en mindre gruppe af patienter med fibromyalgi har gavn af medicinen. Gennem det seneste årti har man kunnet observere et stigende forbrug af lav dosis naltrexone (LDN) til behandling af fibromyalgi. Der findes kun få små studier der har undersøgt effekten af 4.5 mg LDN på fibromyalgi. Disse studier har vist lovende resultater, men kvaliteten af studierne har generelt været lav og det øger risikoen for at effekten bliver overvurderet. Ingen studier har undersøgt effekten af andre doser end 4.5 mg LDN til behandling af fibromyalgi. De overordnede formål med denne afhandling er 1) at undersøge træfsikkerheden af nye spørgeskema baserede diagnose kriterier for fibromyalgi blandt patienter henvist til et Smertecenter, og 2) at undersøge effekten af LDN til behandling af patienter med fibromyalgi. Afhandlingen omfatter et studie omkring diagnostisk træfsikkerhed (**Studie I**) og to lægemiddelforsøg med LDN, et dosis forsøg (**Studie II)** and et lodtrækningsforsøg (**Studie III**).

Studie I undersøgte træfsikkerheden af de såkaldte 2016 kriterier for fibromyalgi blandt patienter med forskellige typer af kroniske smerter henvist til et Smertecenter. Der er ikke tidligere lavet lignende studier. Studiet viste en høj træfsikkerhed af 2016 diagnose kriterierne og understøtter, at de kan bruges på Smertecentrene til at identificere patienter med fibromyalgi både til kliniske og forskningsmæssige formål.

Studie II havde til formål at kvalificere en passende test dosis til det efterfølgende lodtrækningsforsøg (**Studie III**). I **Studie II** blev LDN doser op til 6 mg afprøvet hos kvinder med fibromyalgi og forsøget pegede på, at 6 mg kunne være en mere effektiv test dosis end 4.5 mg. Det blev derfor besluttet at teste effekten af 6 mg LDN sammenlignet med placebo hos kvinder med fibromyalgi i **Studie III**. Resultaterne fra lodtrækningsstudiet viste ingen generelt god smertestillende effekt af LDN sammenlignet med placebo. Derimod pegede resultaterne på at nogle

fibromyalgi patienter måske kan have en klinisk relevant smertestillende effekt af behandling med LDN. Effekten på en række andre fibromyalgi symptomer blev også undersøgt, og her pegede resultaterne på en mulig effekt på hukommelsesproblemer. Studiet afslørede ingen problemer med sikkerheden i forhold til brugen af LDN. **Studie III**, som er det til dato største forsøg af sin slags, blev udført efter de højeste standarder for udførelse og rapportering af kliniske lægemiddelforsøg, og bidrager væsentligt til vores viden om effekten af LDN til behandling af fibromyalgi.

De overordnede konklusioner på denne afhandling er: 1) De spørgeskema baserede 2016 kriterier kan anvendes af Smertecentre til at identificere patienter med fibromyalgi, både til kliniske og forskningsmæssige formål. 2) Der ser ikke ud til at være en generel effekt af LDN på smerter hos patienter med fibromyalgi, men der er sandsynligvis nogen patienter med fibromyalgi, som har en klinisk relevant smertelindrende effekt. Der er fortsat brug for flere lodtrækningsforsøg i fremtiden, før man kan konkludere noget endeligt om den smertestillende effekt af LDN på fibromyalgi.

Abbreviations

ICD	International Classification of Diseases
NSAIDs	Non-steroidal anti-inflammatory drugs
ACR	American College of Rheumatology
FSQ	Fibromyalgia Survey Questionnaire
LDN	Low dose naltrexone
CNS	Central nervous system
TSP	Temporal summation of pain
QST	Quantitative sensory testing
fMRI	Functional magnetic resonance imaging
CPM	Conditioned pain modulation
GABA	γ-aminobutyric acid
CSF	Cerebrospinal fluid
MOR	Mu opioid receptors
WPI	Widespread pain index
SSS	Symptom severity score
CWP	Chronic widespread pain
FIQ	Fibromyalgia Impact Questionnaire
FIQR	Fibromyalgia Impact Questionnaire revised
NRS	Numeric rating scale
SIQR	Symptom Impact Questionnaire revised
EULAR	The original European League Against Rheumatism
TLR	Toll-like receptor
TCA	Tricyclic antidepressant the original
RCT	Randomised controlled trial
SMD	Standard mean difference
RR	Risk ratio
CI	Confidence interval
SNRI	Serotonin noradrenaline reuptake inhibitor
VAS	Visual analoug scale
OMERACT	Outcome Measures in Rheumatological Clinical Trials
PGI-C	Patient global impression of change
ED50	Median effective dose in 50%
ED95	Median effective dose in 95%
PGI-I	Patient Global Impression of Improvement
ISI	Insomnia severity index
ITT	Intention-to-treat
PP	Per Protocol
NNT	Number needed to treat
NNH	Number needed to harm
MID	Minimal clinical difference

Table of Contents

Chapter 1. Introduction

With the 11th revision of the International Classification of Diseases (ICD) in 2018, chronic pain was recognised as a disease in its own right (1). Fibromyalgia represents a well-described subgroup of patients with chronic pain, classified as a chronic primary pain disorder according to ICD-11 (2). Although the exact underlying aetiology is not fully understood, fibromyalgia is considered a central sensitivity syndrome, with amplification of sensory and nociceptive signals leading to diffuse pain and tenderness combined with somatic symptoms and environmental hypersensitivity (3). Mechanisms underlying chronic pain can be nociceptive, neuropathic, nociplastic, or a mix of these (4). Nociplastic pain refers to a state of augmented pain processing with fibromyalgia as a proposed prototype (5, 6). Fibromyalgia has a prevalence of about 2% in the general population, is strongly associated with disability and often has a high impact on activities of daily life, working capacity, and social participation (7, 8).

In Denmark, patients with chronic pain, including fibromyalgia, can be referred for treatment at a specialised pain care centre, where different treatment elements are integrated in a personalised approach (9). Pharmacological treatments aim to reduce pain intensity and/or increase daily function and should preferably be mechanism-based instead of symptom-based (10). Analgesics that are effective for nociceptive pain, such as acetaminophen, non-steroidal anti-inflammatory drugs (NSAIDs) and opioids, have proven ineffective for the treatment of fibromyalgia (11). Current guideline-recommended pharmacological treatments for fibromyalgia are centrally-acting drugs that act by lowering excitatory neurotransmitters (e.g. pregabalin) or increasing the release of inhibitory neurotransmitters (e.g. duloxetine) (12). Thus, fibromyalgia identification in specialised pain care settings can be valuable for guiding the pharmacological treatment approach.

Rheumatologists have traditionally diagnosed fibromyalgia, and current criteria for diagnosis and classification have been developed and initially validated in rheumatological populations. The widely used American College of Rheumatology (ACR) 1990 tender point based criteria (13) have been criticised as being too difficult to use and for capturing an extreme end of a pain and tenderness spectrum, whereas other essential fibromyalgia symptoms are disregarded (14, 15). Thus, ACR endorsed new symptom-based diagnostic criteria in 2010 based on a survey combining a widespread pain index and a symptom severity score (16). A self-administration version, named the Fibromyalgia Survey Questionnaire (FSQ), has been developed (16). In the initial version, the ACR2010 criteria could be satisfied with only three pain sites if combined with a high symptom

score. These criteria performed well in rheumatological populations, but when subsequently applied to a chronic pain population, they were shown to misclassify a very high percentage of patients with regional pain syndromes as having fibromyalgia (17). As a result, a revision was made in 2016, adding a generalised pain criterion, requiring pain in 4 of 5 body sites (18). The 2016 fibromyalgia criteria have shown good discriminative power in rheumatological populations (19). These new symptom-based criteria could be potentially valuable as an easy-to-apply method for the identification of fibromyalgia in chronic pain populations.

Many patients with fibromyalgia fail to respond to or have intolerable side effects from guideline-recommended pharmacological treatments (20, 21), and new effective treatment options are highly warranted. In 2014, low dose naltrexone (LDN) was proposed as a new promising treatment for fibromyalgia and other chronic pain syndromes (22), and a rising utilisation was subsequently observed (23). The evidence for a pain effect has only been supported by a few small studies testing the efficacy of LDN versus placebo in women with fibromyalgia (24, 25). These studies had several methodological weaknesses, with a risk of overestimating the effect (26). Furthermore, issues regarding dose-response have not been studied, and a similar dose of 4.5 mg was used in all the preceding LDN/fibromyalgia trials.

In summary, pain specialists are in need of well-established diagnostic criteria for identifying fibromyalgia to support a mechanism-based pharmacological treatment strategy, and the introduction of new treatments must be based on evidence regarding dosing, efficacy, and safety.

1.1. Aims of the thesis

The overall aim of the current thesis was to improve current knowledge regarding diagnosis of fibromyalgia in specialised pain care settings and treatment efficacy of LDN for fibromyalgia.

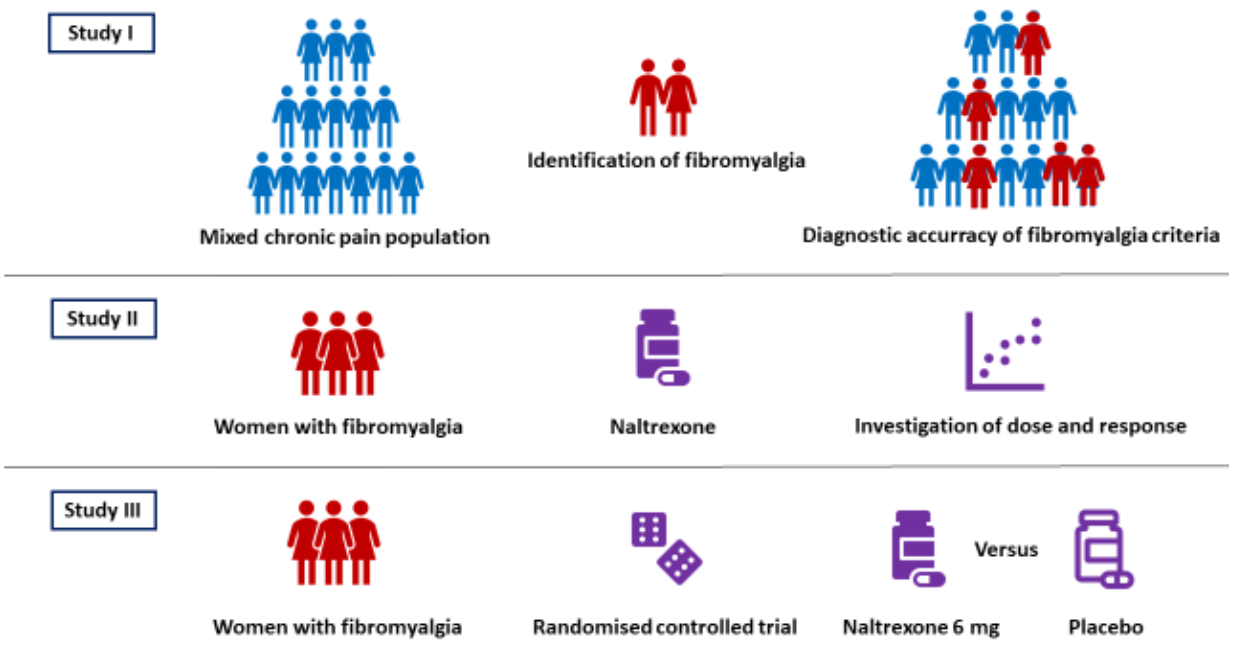
Aims at the study level were:

Study I: To investigate the diagnostic accuracy of the 2016 diagnostic criteria for fibromyalgia in a population of patients with mixed chronic pain referred to specialised pain care.

Study II: To explore dose-response relationships for naltrexone in a low dose range on the improvement of global impression and pain in women with fibromyalgia, using the “up-and-down” method.

Study III: To investigate if treatment with LDN for 12 weeks had a superior effect on pain in women with fibromyalgia compared to placebo. The secondary aims were to investigate pain-responder indices and effects on other symptoms, daily functioning, total impact, health-related quality of life, and global impression following treatment with LDN for 12 weeks.

1.2. Overview of methods and materials used in the studies included in the thesis



	Study I	Study II	Study III
Paper	I	II	III + IV
Population	**A cross-sectional sample of patients referred to two Danish specialised pain care centres N=215**	**Women with fibromyalgia who had completed treatment at a Danish specialised pain care centre N=27**	**Women with fibromyalgia recruited from all over Denmark N=99**
Design	**A prospective diagnostic accuracy study**	**A prospective dose-response study**	**A randomised, double-blinded, placebo-controlled trial**
Method	**The performance of the 2016 diagnostic criteria was assessed using a clinical diagnosis of fibromyalgia based on the expert opinion of a pain specialist as the reference standard.**	**Naltrexone in the dose range between 0.75 mg and 6 mg, with dose intervals of 0.75 mg, were tested using the "up-and-down" method.**	**Participants were allocated 1:1 to treatment with Naltrexone 6 mg once daily or an identically appearing placebo tablet. The treatment period was 12 weeks.**

Chapter 2. Background on fibromyalgia

The clinical phenotype associated with fibromyalgia has been described for centuries and labelled with different descriptive terms such as "non-articular rheumatism," "muscular rheumatism," or "fibrositis" (27). When consensus classification criteria, endorsed by the ACR, was published in 1990, the term "fibromyalgia" was adopted. Fibromyalgia was subsequently incorporated as a diagnosis (as a subcategory of unspecified rheumatism) in the World Health Organization's 10th revision of the ICD from January 1st, 1993 (28).

With the recent 11th revision of ICD in 2018, fibromyalgia is now classified as a chronic primary pain disorder (2). Chronic primary pain is a top-level diagnosis, with chronic widespread pain and fibromyalgia representing second and third-level diagnoses, respectively (Figure 2-1).

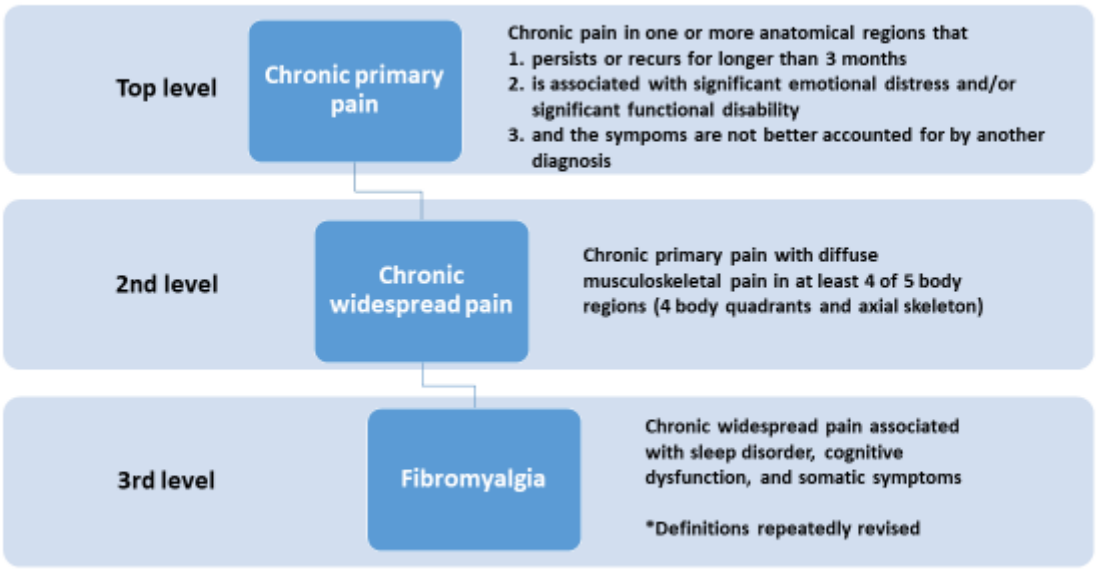


Figure 2-1. Scematic overview of the hierachy of pain diagnoses according to ICD-11, with fibromyalgia as a third level diagnose

2.1. The clinical features of fibromyalgia

The clinical characteristics of fibromyalgia (Figure 2-2) can be attributed to a generalised state of central sensitisation, with amplification of nociceptive, non-nociceptive and environmental stimuli (3). The dominant feature of fibromyalgia is diffuse pain and tenderness perceived from

musculoskeletal tissues, where no structural changes or inflammation can be identified (11). A reduced threshold to mechanical and noxious stimuli clinically presents as diffuse allodynia or hyperalgesia to pressure, touch, heat or cold. Central symptoms such as disturbed sleep, fatigue, and dyscognition are present at a moderate to severe level. The amplification of non-nociceptive body signals can cause symptoms from all organ systems, e.g. sensory hypersensitivity, nausea, dyspepsia, dysuria, hyperactive bladder, paraesthesias, balance problems, dizziness, tinnitus, blurred vision, shortness of breath, palpitations, dysmenorrhea, itching, and subjective fevers. Pain is often aggravated by activity, emotional distress, and environmental factors (e.g. weather changes) (3). Earlier adverse life events might contribute to a cognitive-emotional sensitisation, which clinically presents as low-stress tolerance, worry/anxiety or dysthymia/depression (29).

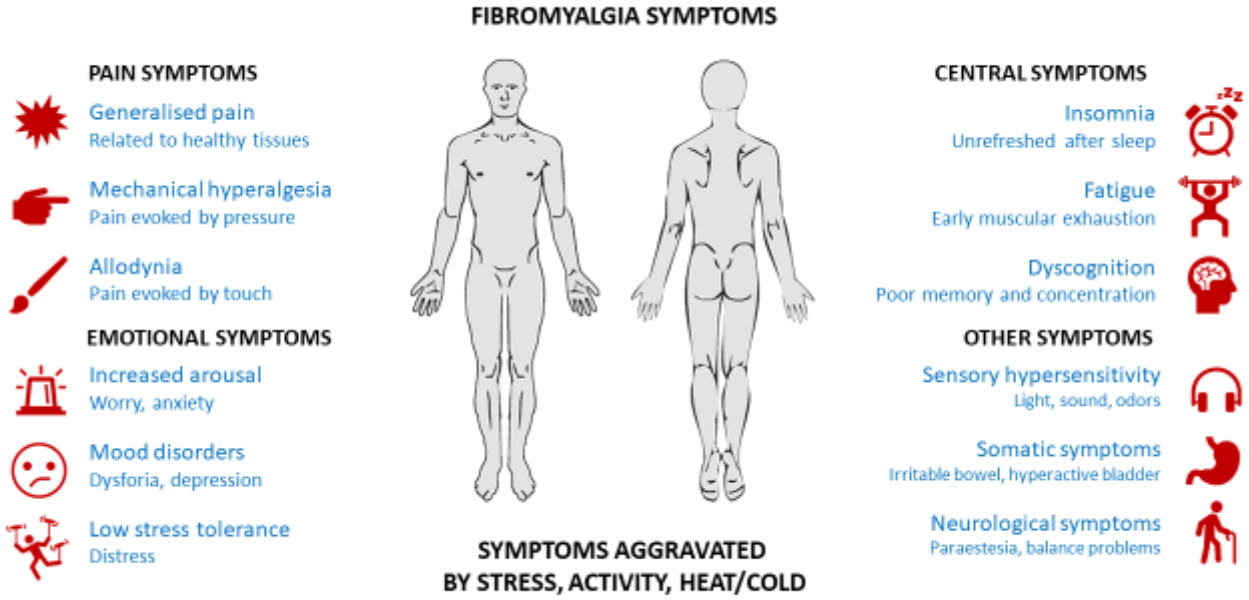


Figure 2-2. The clinical characteristics of fibromyalgia

2.2. Prevalence in general and patient populations and gender distribution

In a 2017 review, the total prevalence of fibromyalgia was estimated to be 1.78% in the general population worldwide (8). However, a very high variability of estimates has been found across many prevalence studies (8, 30). These differences reflect a high heterogeneity in the criteria used to identify fibromyalgia (Table 2-1).

Table 2-1. Overview of criteria used for the identification of fibromyalgia in prevalence studies.

Criteria used in different prevalence studies
ACR1990 criteria
ACR2010 criteria
Fibromyalgia Impact Questionnaire (FIQ)
London Fibromyalgia Epidemiology Study Screening Questionnaire (LFESSQ)
Community Oriented Program for the Control of Rheumatic Diseases (COPCORD)
ICD-10
Self-reported by patients
Clinical diagnosis based on expert opinion

Which criteria are used to validate a fibromyalgia diagnosis can also greatly influence the estimated gender distribution (31). Prevalence studies in patient populations using a clinically verified diagnosis of fibromyalgia typically estimate that more than 90% of cases are women (32). When applying symptom-based criteria in population-based studies, the gender distribution is more even, with about 60% of cases being women (31). These differences might reflect that the composition of patient samples are biased by several factors, e.g., referral patterns, patient and physician beliefs, etc. In contrast, symptom-based criteria applied to a general population sample will identify all individuals fulfilling the criteria, including non-diagnosed individuals (31).

Secondary-concomitant fibromyalgia is found in up to 30% of patients suffering from painful inflammatory or autoimmune diseases (e.g. rheumatoid arthritis, systemic lupus erythematosus, or ankylosing spondylitis), contributing to increased morbidity (33, 34). In populations with central sensitivity pain syndromes such as migraine, tension-type headache, temporomandibular disorder, irritable bowel disease and interstitial cystitis, up to 40% have co-morbid fibromyalgia (34). A higher prevalence of fibromyalgia has also been found in patients with diabetes, heart, renal, pulmonary, psychiatric, and other diseases (35). An increasing number of comorbidities increases the risk of concomitant fibromyalgia, with a prevalence of 55% in patients with four or more comorbidities (35).

2.3. Pathophysiological mechanisms

Much evidence supports that fibromyalgia is associated with abnormal pain processing, with signs of peripheral and central sensitisation, altered modulation of pain signals with imbalances in excitatory and inhibitory neurotransmitters, and altered functional connectivity in pain-processing

brain areas (36). Both genetic factors and "psychological sensitisation" seem to predispose some individuals to an augmented pain perception (29). More recent evidence has revealed an association with increased neuro-inflammatory activity and increased levels of proinflammatory cytokines in blood and cerebrospinal fluid (29, 36). Furthermore, alterations in neuroendocrine, endogenous opioid, and dopamine activity have been demonstrated, contributing to a highly complex interplay of many pathophysiological mechanisms (29, 36).

2.3.1. Peripheral sensitisation

Studies investigating whether changes in peripheral tissues were accountable for fibromyalgia symptoms have failed to demonstrate any pathology in muscle biopsies or metabolism (37-39). Several studies have shown pathology in small nerve fibres (e.g. unmyelinated C fibres and low threshold mechanoreceptors mediating pain, heat and cold sensations), and reduced intra-epidermal nerve fibre density has been estimated to have a prevalence of about 50% in patients diagnosed with fibromyalgia (40). Furthermore, spontaneous activity in silent C fibres has also been demonstrated (41).

2.3.2. Central sensitisation

Central sensitisation refers to a state of higher responsiveness and increased signalling in central nervous system (CNS) pathways (3). Studies of temporal summation of pain (TSP), a quantitative sensory testing (QST) paradigm developed as a human model equivalent to the wind-up phenomena demonstrated in animals, show that patients with fibromyalgia have a significantly larger response to repeated noxious stimuli than healthy controls (42).

Several brain imaging studies using functional magnetic resonance imaging (fMRI) techniques have not demonstrated increased brain activation during pain stimulation, suggesting that the CNS amplification of pain signals occurs at subcortical levels (43). However, other studies have shown increased connectivity in pain processing networks during resting state in patients with fibromyalgia (43).

2.3.3. Descending pain modulation

Studies of conditioned pain modulation (CPM), another QST paradigm developed as an experimental method of assessing the inhibitory capacity within the CNS, show that patients with fibromyalgia have a smaller inhibitory response to painful stimuli than healthy controls (42). Only a

few fMRI and electroencephalography (EEG) studies have investigated the descending pain inhibitory pathways, with some evidence supporting reduced activity (43, 44).

2.3.4. Neurotransmitters

At the level of the neuronal synapses, the balance between the excitatory neurotransmitter glutamate and the inhibitory neurotransmitter γ-aminobutyric acid (GABA) determines whether a postsynaptic action potential (e.g. pain signal) is being generated or not. Glutamate is synthesised from Gluatamine, which can be converted into GABA in the so-called Glutamate/GABA-glutamine cycle (45). Therefore, the levels of these two essential neurotransmitters are closely connected and are being recycled in collaboration with astrocytes.

Several studies have shown that patients with fibromyalgia have higher levels of the excitatory neurotransmitters glutamate and substance P in the cerebrospinal fluid (CSF) and the brain (visualised with proton magnetic resonance spectroscopy (^{1}H-MRS)) (46, 47). There is also some evidence that levels of GABA, the main inhibitory neurotransmitter, are reduced in the brain of patients with fibromyalgia (48).

2.3.5. Imbalance in the endogenous opioid system

There is some evidence pointing to deficient endogenous opioid analgesia in patients with fibromyalgia. This endorphin system imbalance has been characterised by decreased availability of mu-opioid receptors (MOR) receptors in some brain areas, which might result from receptor downregulation due to excess brain opioid peptides (49, 50).

2.3.6. Neuroinflammation

Brain glial cells, such as microglia and astrocytes, are neuroimmune cells that are thought to be closely involved in chronic pain pathogenesis (51). Activated glial cells release cytokines that sensitise peripheral and central sensory neurons and thereby augment the activity in pain circuits. Some of the cytokines released from activated glial cells, such as interleukin-6 and tumour necrosis factor, have been found to be significantly elevated in blood and CSF in patients with fibromyalgia compared to healthy controls (52). These cytokines are not specific for glial activity and can also be released from peripheral cell types as macrophages. However, a recent study using positron emission tomography (PET) has demonstrated increased activity of microglia, but not astrocytes, in the brain of patients with fibromyalgia compared to healthy controls (53).

2.4 Diagnosing fibromyalgia

None of the biochemical changes found in body fluids or aberrations demonstrated through experimental tests or brain imaging techniques have proven specific for fibromyalgia, and no gold standard exists to validate the diagnosis. Thus, fibromyalgia continues to be a clinical diagnosis based on the presence of a cluster of symptoms that significantly impact daily function.

Several sets of classification and diagnostic criteria have been developed to assist in the identification of patients with fibromyalgia in research and clinical practice. Generally, classification criteria are designed for research to identify well-defined groups of patients with high specificity. In contrast, diagnostic criteria aim to capture the whole spectrum of the disease and require high sensitivity (54).

2.4.1. Consensus classification criteria for fibromyalgia

The first set of criteria (Smythe's criteria) to identify fibrositis/fibromyalgia was published in 1977, introducing a range of non-pain symptoms and tender points as diagnostic items (55). During the following decade, several diagnostic criteria (e.g. Bennett's and Yunus' criteria) were proposed by different research groups (56, 57), with disagreements mainly concerning whether criteria should emphasise multiple symptoms or multiple tender points (58). A committee with participants from 16 rheumatological centres in the United States of America and Canada was established in 1986 to develop consensus classification criteria. All previous criteria and single items were tested (using robust methodology) in a large patient population sample from the involved rheumatological settings (13). The consensus criteria were published in 1990, endorsed by the ACR, and are usually referred to as the ACR1990 criteria (13).

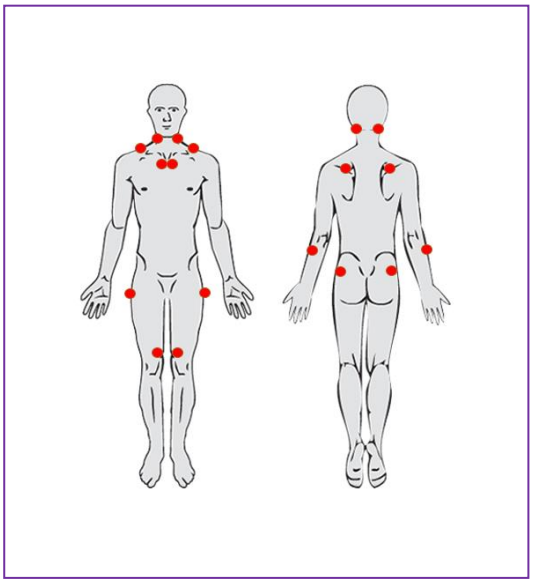
2.4.2. The American College of Rheumatology 1990 classification criteria

Among all the single items and combinations tested in the development study, the combination of widespread pain and 11 of 18 tender points had the highest discriminative power. Thus, the ACR1990 classification criteria required 1) a history of widespread pain over three months and 2) pain in 11 of 18 tender points on digital palpation with approximately 4 kg. The criteria did not exclude the presence of another disorder and were shown to work equally well for diagnosing secondary-concomitant fibromyalgia.

Widespread pain was defined as pain on the right side of the body, pain on the left side of the body, pain above the waist, pain below the waist, and pain in the axial skeleton (13). This ACR1990

widespread pain definition has been subject to different interpretations (59); however, in the original article, it is stated that three pain sites (e.g. right shoulder, left buttock, and thoracic spine) qualify as widespread pain (13).

Initially, 24 tender points were defined, and a scoring system for grading the severity of tenderness was defined as 0 = no pain, 1 = mild pain (vocalised), 2 = moderate pain (grimace or flinch), and 3 = severe pain (withdrawal). By analysing the discriminatory power of each tender point, the number was reduced to 18, representing nine bilateral sites (Figure 2-3) (13). A site was defined as a tender point in case of mild (or greater) pain, on digital examination with the pulp of the thumb, standardised as a 4 kg pressure.



Occiput	At the suboccipital muscle insertions
Low cervical	At the anterior aspects of the intertransverse spaces at C5-C7
Trapezius	At the midpoint of the upper border
Supraspinatus	At origins, above the scapula spine near the medial border
Second rib	At the second costochondral junctions, lateral to the junctions on upper surfaces
Lateral epicondyle	2 cm distal to the epicondyles
Gluteal	In upper outer quadrants of buttocks in anterior fold of muscle
Greater trochanter	Posterior to the trochanteric prominence
Knee	At the medial fat pad proximal to the joint line

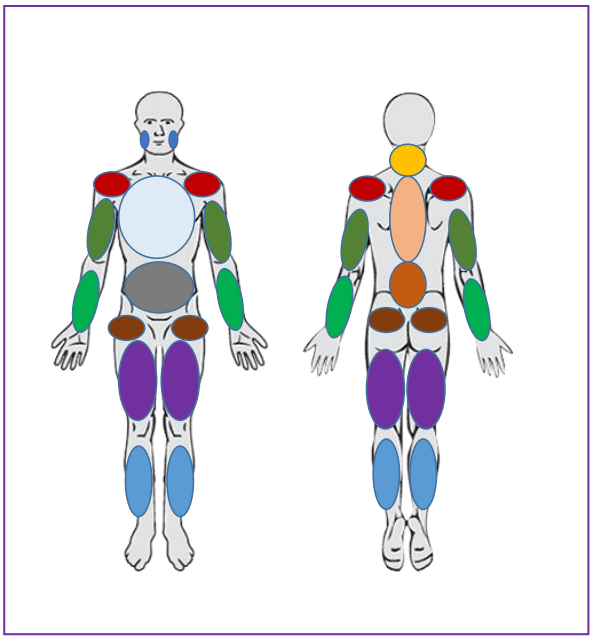
Figure 2-3. Tender point defintions according to the American College of Rheumatology classification criteria

2.4.3. The 2010 diagnostic criteria

In the following decades, the ACR1990 criteria were intensely debated (60). A practical argument against the criteria was that tender point examination required training and was difficult to use outside speciality clinics (14, 15). Others argued that patients meeting the criteria represented an extreme end of a pain and tenderness spectrum. In contrast, other significant fibromyalgia symptoms were disregarded, failing to "capture the essence of the fibromyalgia syndrome" (61, 62).

A new multicenter study was designed to develop new non-tender point diagnostic criteria, to accommodate the need for more feasible diagnostic criteria suitable for use outside speciality clinics and the wish for a symptom severity scale for longitudinal follow-up. These criteria, published in 2010 and endorsed by the ACR as provisional, were not meant to replace the ACR1990 criteria but to provide an alternative (16). The ACR2010 criteria are based on a survey consisting of a Widespread Pain Index (WPI) and a Symptom Severity Score (SSS).

The WPI was defined based on the presence of pain in 19 non-articular sites (Figure 2-4), giving a continuous measure of widespread pain ranging from 0-19 (16). WPI was highly correlated to tender point count, and a WPI ≥ 7 classified 83,6% of cases correctly when using the ACR1990 criteria as the reference standard (16).



Shoulder	Right/left
Hip	Right/left
Upper arm	Right/left
Lower arm	Right/left
Upper leg	Right/left
Lower leg	Right/left
Jaw	Right/left
Axial	Low back Upper back Neck
Thorax	
Abdomen	

Figure 2-4. The Widespread Pain Index Definition of Pain Sites

The SSS was constructed based on an analysis of correlations for non-pain variables. The symptom variables that correlated best with tender point count were fatigue, cognition, waking unrefreshed, and somatic symptoms. The extent of somatic symptoms was graded using a reference list of 41 symptoms (0 = no symptoms, 1 = few symptoms, 2 = a moderate number of symptoms, 3 = a great deal of symptoms). The other three symptoms were graded regarding severity the last week, using a 4-point Likert scale (0 = no problem, 1 = slight or mild problems, 2 = moderate, considerate

problems, 3 = severe, pervasive, continuous, life-disturbing problems). The SSS was created by summing these four symptom variables' 0-3 scores, giving a 0-12 range (16).

The 2010 diagnostic criteria are defined as a combination of 1) WPI ≥ 7 and SSS ≥ 5 *or* 2) WPI ≥ 3 and SSS ≥ 9. It is required that the symptoms should have been present at a similar level for at least three months and that other disorders could not sufficiently explain the pain (16).

2.4.4. The Fibromyalgia Survey Questionnaire

The original survey required evaluation by a physician, and a modification was developed for self-administration in 2011 (63). The modification mainly consisted of the substitution of the somatic symptom item. Instead of grading the number of somatic symptoms, the prevalence (yes/no) of three symptoms (headache, pain or cramps in the lower abdomen, and depression) during six months was measured. The self-administration survey was named the Fibromyalgia Survey Questionnaire (FSQ) and was initially recommended for use in research only (64).

2.4.5. The 2016 diagnostic criteria

The ACR2010 criteria were tested in subsequent studies, showing about 85% agreement with the ACR1990 criteria (17). When investigating populations from rheumatological settings, which included patients diagnosed with fibromyalgia by a physician, the ACR2010 criteria were satisfied in most FM cases (65). However, problems arose when applying the criteria to a mixed chronic pain population. In a study by Egloff et al., a population of 300 patients with different pain syndromes recruited sequentially from a specialised pain care setting showed that a very high percentage of patients with regional pain syndromes were misclassified as fibromyalgia (17). Among 5011 patients from the National Data Bank of Rheumatic Diseases fulfilling the ACR2010 criteria, 93,8% reported pain in all five regions (4 quadrant plus axial), and 98.8% reported pain in at least four regions when using the five region pain definition from the ACR1990 criteria (66). As a result of these findings, a generalised criterion was added to a 2016 revision of the diagnostic criteria (18).

The 2016 criteria are based on the FSQ, which the patient or a physician can fill out, but only a physician can evaluate if a patient should be diagnosed with fibromyalgia (18). The 2016 diagnostic criteria require the combination of the following:

1) Generalised pain, defined as pain in at least 4 of 5 regions
2) Symptoms have been present at a similar level for at least three months
3) WPI ≥ 7 and SSS score ≥ 5 *or* WPI of 4–6 and SSS score ≥ 9

2.4.6. The performance of the 2016 criteria in chronic pain populations

Scrutinising the literature, four studies (Table 2-2) investigating the diagnostic accuracy of the 2016 criteria were discovered (67-70). Only one of these studies investigated the performance of the 2016 criteria in a specialised pain care setting, showing low sensitivity and specificity (69). However, in this study, only patients with chronic widespread pain (CWP) were invited to participate, and fibromyalgia cases were defined as participants with CWP who fulfilled the ACR1990 criteria.

Table 2-2. Studies investigating the diagnostic accuracy of the 2016 criteria with estimates of the sensitivity and specificity

Country (reference)	Recruited from	Population	Reference standard	Sensitivity	Specificity
Korea (67)	**Rheumatological setting**	**FM versus Inflammatory rheumatic diseases**	**Clinical diagnose**	**93.1%**	**90.7%**
Norway (68)	**Patient associations**	**FM versus Mixed general population**	**ACR1990 criteria**	**88.8%**	**77.5%**
Italy (69)	**Rheumatological setting**	**Non-inflammatory rheumatic diseases**	**Clinical diagnose**	**78.0%**	**90.5%**
India (70)	**Specialised pain care setting**	**Chronic widespread pain**	**ACR1990 criteria**	**71%**	**60%**

2.5. Assessment and monitoring of fibromyalgia

The developers of the ACR2010 criteria have suggested that the summed score of WPI and SSS, named the polysymptomatic distress scale (PDS), can be used to measure disease severity and for longitudinal follow-up (71). However, the most widely used tool for assessing and monitoring disease severity is the fibromyalgia impact questionnaire (FIQ), a validated self-report disease-specific questionnaire covering the spectrum of problems associated with fibromyalgia (72). A revised version (FIQR) was validated and published in 2009, with improvements in wording and changes in constructs that could be a potential source of bias in the original version, such as gender bias and ethnicity bias (73). The FIQ/FIQR is sensitive to changes in fibromyalgia symptomatology and has good discriminatory power between fibromyalgia and chronic pain of other origins (72, 73). Thus, the questionnaire has also been recommended as an outcome measure for use in research (74).

2.5.1. The Fibromyalgia Impact Questionnaire-Revised (FIQR)

The FIQR contains domains of function (range 0-90), impact (range 0-20), and symptom severity (range 0-100), with lower scores indicating lower severity. The total score (range 0-100) is obtained by summing the impact score with 1/3 of the function score and 1/2 of the symptom score. The function domain evaluates daily functions by rating how difficult it is to perform a list of 9 everyday activities over the previous seven days on a 0-10 numeric rating scale (NRS). The symptom domain evaluates the severity of pain and nine non-pain fibromyalgia symptoms on a 0-10 point NRS. The impact domain consists of two questions assessing fibromyalgia's overall physical and emotional interference during the previous seven days on a 0-10 NRS.

The Symptom Impact Questionnaire-Revised (SIQR) is identical to FIQR, except the word "fibromyalgia" is replaced with "symptoms" or "medical problems" throughout the questionnaire (75, 76). The SIQR is used to assess the classical features of fibromyalgia in non-diagnosed patient populations.

2.5.2. The FIQR Danish version

We have translated the FIQR in Danish, according to standardised guidelines, comprising the following steps: forward translation, backward translation, panel discussion, pilot testing, and final version. Our Danish version of the FIQR is included in **Appendix A**.

We also validated our Danish version and found excellent test-retest reliability and internal consistency using the Intraclass Correlation Coefficient and Cronbach's alpha coefficient, respectively. Correlations with the Hospital Anxiety and Depression Scale domains and the 36-Item Short Form Health Survey domains were analysed using Spearman's rho correlation coefficient. We found good correlations between FIQR depression and anxiety and the Hospital Anxiety and Depression Scale. Correlations with 36-item Short Form Health Survey domains ranged from fair for FIQR energy, good for FIQR pain, and very good for FIQR function. Results from the validation study will be published in a peer-reviewed journal (manuscript currently in review).

Chapter 3. Identification of fibromyalgia in a specialised pain care setting

The aim of **Study I** was to investigate if the 2016 diagnostic criteria could identify fibromyalgia correctly when applied to a population of patients with mixed chronic pain syndromes referred to a specialised pain care setting.

To do this, we designed a prospective diagnostic accuracy study adhering to the Standards for Reporting Diagnostic Accuracy Studies (STARD) guidelines from 2015 (77). The performance of the 2016 criteria was assessed using a clinical diagnosis of fibromyalgia as the reference standard based on the opinion of a skilled pain specialist.

The study design, main findings, methodological considerations and implications of the findings are briefly reviewed below. Please refer to **Paper I** for further details.

3.1. Methods and materials

The participants for **Study I** were recruited among patients who started their treatment at two specialised pain care centres in The Region of Southern Denmark (Odense and Middelfart) on December 10, 2017, and for two consecutive years (Figure 3-1).

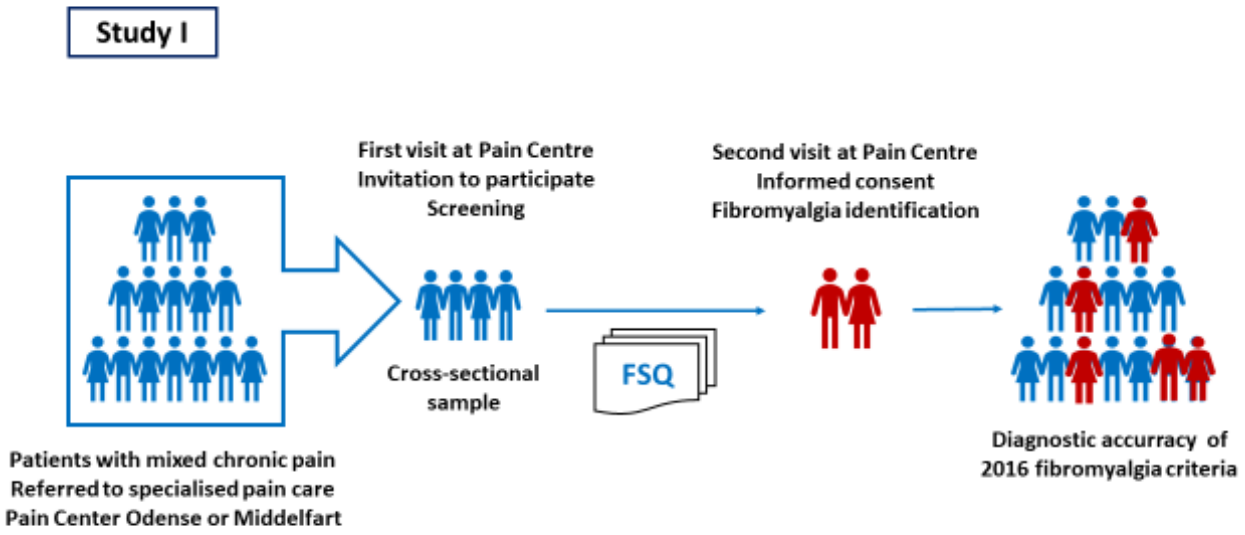


Figure 3-1. Scematic overview of the study flow in study I

A cross-sectional sample of all the patients referred in the period was obtained, as only patients who had their first appointment with one specific consultant at each pain centre (KDB or NK) were invited to participate. Inclusion was terminated in August 2019 in Middelfart because the investigating consultant (NK) changed position. To be eligible, patients had to be 18 years or older, and they were required to read and understand Danish. Patients who could not complete questionnaires due to poor bodily or mental health were excluded.

3.1.1. The FSQ Danish version

As part of study I, the FSQ was translated to Danish according to standardised guidelines, as previously described. No cultural adaptations were needed. The Danish version of the FSQ is available in **Appendix B**.

3.1.2. The reference standard

Participants who consented and were eligible were evaluated by one of the two investigators to establish if criteria were met for a clinical diagnosis of fibromyalgia. The two consultants were experienced pain specialists and had received training in diagnosing fibromyalgia from a senior rheumatologist and fibromyalgia expert before initiating the study. The diagnosis was based on a complete medical history, a pain interview, pain drawings, and a tenderpoint examination. The minimum requirements for a clinical diagnosis of fibromyalgia were patient-reported pain in all four body quadrants plus axially for a minimum of 3 months and the presence of at least 8 of 18 positive tender points as defined by the ACR1990 criteria (13). A handheld spring-based pressure algometer with a probe area of 1 cm2 was used instead of digital palpation to reduce interrater variability. Each site was examined by slowly increasing the pressure to a maximum of 4 kg/cm^2.

3.1.3. The test under evaluation

All participants filled out a paper version of the FSQ in Danish before the clinical evaluation. The investigators were blinded regarding the FSQ score during the interview and tender point examination. After the diagnostic evaluation, the FSQ was scored to determine whether the 2016 criteria were fulfilled.

3.1.4. Other assessments

Demographic data about age, gender, body mass index, work status, medication, and comorbidity were collected. Based on information from the patients and the medical file, information was obtained on whether the patients had been diagnosed with fibromyalgia before. To evaluate the

severity of the pain and accompanying symptoms, the participants filled out the symptom domain of the SIQR.

3.1.5. Statistics

A contingency table was made showing the performance characteristics of the 2016 diagnostic criteria based on the patient version of the FSQ with a clinical diagnosis as the reference standard. From this table, data were derived to calculate sensitivity and specificity. Furthermore, the positive and negative post-test probability were estimated.

3.2. Main findings of Study I

We estimated the prevalence of fibromyalgia in the present population of patients with mixed chronic pain syndromes referred to a specialised pain care setting to be:

1. 45% based on a clinical diagnosis by the opinion of a pain expert
2. 50% based on the 2016 criteria
3. 41% based on the ACR1990 criteria
4. 19% based on self-report

Gender distributions:

1. Of the total population, 74% were women
2. Among women, 55% were identified as having fibromyalgia based on a clinical evaluation
3. Among men, 15% were identified as having fibromyalgia based on a clinical evaluation
4. In this chronic pain population sample, the relative male-to-female ratio was 1:3.7

The performance characteristics of the 2016 criteria using a clinical diagnosis of fibromyalgia by expert opinion as the reference standard showed:

1. Sensitivity of 89%
2. Specificity of 82%
3. Positive post-test probability of 79%
4. Negative post-test probability of 10%

Patients diagnosed with fibromyalgia in this study had a significantly higher symptom burden assessed by the level of the ten SIQR symptom items compared to chronic pain patients without a fibromyalgia diagnosis. When looking at patients diagnosed with fibromyalgia in the study, no

differences were observed between patients with or without a previously established diagnosis in any of the measured clinical characteristics.

3.3. Methodological considerations

The lack of a gold standard for diagnosing fibromyalgia represents a major challenge when testing the diagnostic accuracy of new criteria. Diagnoses based on self-report or ICD-10 coding will be biased by referral patterns and beliefs of patients and physicians. Furthermore, current status cannot be established reliably by a past diagnosis. Using classification criteria, such as the ACR1990 criteria, as the reference standard will predictably bias the results, as the measurement errors of these criteria will be transferred. Thus, a clinical diagnosis of fibromyalgia made by an expert is presumed to be the most accurate method of establishing a reference standard, even though expert opinions might vary. Studies with a valid recruitment method, as serial or all patients, are less likely to be biased than convenience samples (18).

The main strength of **Study I** was the prospective design, where all patients included were diagnosed simultaneously with applying the tested criteria. Only two pain experts (who had received simultaneous training) diagnosed all patients, aiming to secure a high agreement of the reference standard. Also, seeking to limit inter-rater variability, a set of minimum requirements was agreed upon. These requirements were based on results from a previous Danish study investigating the association between tender point count and symptom burden in patients with CWP, showing a shift in symptom burden at 8/18 tender points (78). In the current study, the prevalence of fibromyalgia was estimated to be 45%, which makes a sample size of 215 sufficient to estimate diagnostic accuracy with acceptable power (79).

The main limitation of **Study I** could be that the same assessor evaluated the reference standard and the test. Measures to reduce bias were applied, as the FSQ survey was completed by the patient at home, the investigators were blinded to the FSQ during the visit, and the FSQ was scored after the diagnosis had been established. However, an independent assessor of the FSQ score would have been preferable. The study population was a cross-sectional sample of patients referred to specialised pain care at two multidisciplinary pain care settings in Southern Denmark. Thus, the prevalence found in our study might not be generalisable to specialised pain care settings in Denmark or other countries.

As part of study I, the FSQ was translated to Danish, but tests to prove the validity of this Danish version were not done. However, a recent validation study from Norway (a country culturally and linguistically similar to Denmark) has demonstrated that a Norwegian version of the FSQ has good psychometric properties based on analyses of test-retest reliability, construct validity, and internal consistency (68).

3.4. Implication of findings

The 2016 criteria were demonstrated to have high sensitivity in the current study population, and the FSQ could easily be applied as a screening tool for assisting in the identification of fibromyalgia among patients with chronic pain referred to specialised pain care. The specificity of the 2016 criteria was acceptable in the present population and supports using the criteria for research purposes.

Chapter 4. Background on pharmacological treatment and low dose naltrexone

Based on guidelines for the management of fibromyalgia from the original European League Against Rheumatism (EULAR), no generally effective pharmacological or non-pharmacological treatments exist, and most recommended treatments have small effect sizes (80). Therapies targeting peripheral pain mechanisms, such as NSAIDs, massage, or chiropractic treatments, have proven ineffective for fibromyalgia pain (81-83), and EULAR recommendations are against such therapies (80). In patients with fibromyalgia, the balance between CNS excitation and inhibition is shifted towards greater excitation (36). Pharmacological agents such as gabapentinoids and antidepressants, recommended for the treatment of neuropathic pain, act by increasing inhibitory neurotransmitters such as serotonin and noradrenalin or by decreasing glutamic (excitatory) activity in the CNS (84). Several studies have investigated the efficacy of these drugs for the treatment of fibromyalgia, showing effect in some patients but with a risk-benefit profile that is less advantageous than for neuropathic pain (12, 84).

Besides the serotonin/noradrenalin pathway, inhibitory pain modulation can also occur via the opioid pathway. There is no evidence regarding the efficacy of strong opioid drugs for the treatment of fibromyalgia, and EULAR guidelines hold a strong recommendation against these drugs due to the high risk of addiction and other serious adverse effects (80). Tramadol, a weak opioid agonist with some inhibitory activity of serotonin/noradrenaline reuptake, has shown some effect on fibromyalgia pain (85). However, recent evidence shows that tramadol has the same risk of tolerance and addiction as strong opioids (86).

At the beginning of this century, anecdotal reports of a beneficial effect on fibromyalgia from treatment with naltrexone, an opioid antagonist, began to emerge on internet pages (87). It was hypothesised that treatment with LDN via a temporary opioid receptor blockade would lead to upregulation of opioid receptors and opioid ligands, with an improved endogenous opioid function impacting both pain and general well-being (87). The discovery that naltrexone also has antagonistic properties at Toll-like receptors (TLR) found in microglia and astrocytes led to another competing hypothesis that LDN impacts pain and central symptoms via an anti-inflammatory pathway (22). Only a few small clinical trials have investigated the potential efficacy of LDN on

fibromyalgia. No official recommendations are currently available regarding the use of LDN for treating fibromyalgia.

4.1. The evidence of recommended pharmacological treatments for fibromyalgia

In the most recent EULAR guidelines for the management of fibromyalgia, a weak recommendation is given for treatment with amitriptyline, duloxetine, pregabalin, cyclobenzaprine (a muscle relaxant), and tramadol (80). The National Danish guidelines for the treatment of chronic widespread pain (revised in 2018, however, currently expired) give a weak recommendation for the use of amitriptyline, duloxetine, gabapentin and pregabalin but a weak recommendation against the use of tramadol (88). Cyclobenzaprine is not marketed for use in Denmark. Instead, other muscle relaxants, such as baclofen and tizanidin, are commonly used off-label in specialised pain care settings. A brief review of the current evidence for the efficacy of amitriptyline, duloxetine, gabapentin and pregabalin for treating fibromyalgia pain is provided below.

4.1.1. Amitriptyline

Amitriptyline is a tricyclic antidepressant (TCA) with serotonergic, noradrenergic, and anti-histaminergic activities, known to have hypnotic effects in low doses (89). Amitriptyline is a highly non-selective drug, and its antinociceptive effects have been suggested to be mediated primarily through agonism at α_2-adrenergic and serotonergic receptors (89). A 2019 Cochrane review investigated the evidence for an effect on fibromyalgia pain based on four randomised placebo-controlled trials (RCTs) with a total of 275 participants (Table 4-1) (90). The review did not include the standard mean difference (SMD) between groups as an outcome. No data were available regarding a 30% pain reduction. A 50% pain reduction or more was found in 36% of subjects treated with amitriptyline compared to 11% treated with placebo, with a risk ratio (RR) = 3.0 (95% confidence interval (CI): 1.7 to 4.9). The number of withdrawals due to adverse events comprised 8% in the amityline group and 9% in the placebo group. However, the review concluded that there is a substantial risk that the treatment effect was overestimated due to the very low quality of the trials included (90).

4.1.2. Duloxetine

Duloxetine is a serotonin noradrenaline reuptake inhibitor (SNRI) with an equally high affinity to serotonin and noradrenaline transporters (91). The simultaneous increase of these two inhibitory neurotransmitters has been suggested to be crucial for mediating an analgesic effect (92). Duloxetine has a low affinity to adrenergic and serotonergic receptors (91). A 2010 review

investigating the effect of duloxetine on fibromyalgia pain analysed data from seven studies with a total of 2642 participants (Table 4-1) and found significantly greater pain relief from treatment with duloxetine compared to placebo (21). However, the effect size was small, with an SMD = -0.26 (95% CI: -0.37 to -0.16). At least 30% pain reduction was achieved in 47.6% versus 35.9% in the duloxetine and placebo groups, respectively, with an RR = 1.31 (95% CI: 1.19 to 1.44). Regarding the 50% responder indices, the proportions for the duloxetine and placebo groups were 33.9% versus 23.0%, with an RR = 1.45 (95% CI; 1.27 to 1.66). Withdrawals due to adverse events comprised 15.5% in the duloxetine group and 9.5% in the placebo group.

4.1.3. Gabapentin and pregabalin

As the names insinuates, gabapentin and pregabalin were originally designed as GABA mimetics but have no effect on GABA receptor activity (93). Instead, they bind to alpha-2-delta receptors on voltage-activated calcium channels located on excitatory neurons. Pregabalin has been demonstrated to reduce the release of glutamate in the spinal cord, insula and amygdala and is thought to reduce pain by attenuating signalling in spinal cord ascending pain pathways and by influencing the activity in emotion-processing brain regions activated by pain (93).

Gabapentin, which holds a strong recommendation for the treatment of neuropathic pain (84), has only been investigated for fibromyalgia pain in one RCT with 150 participants (Table 4-1) (94). This RCT had a parallel 1:1 allocation design, and participants were treated with gabapentin (or an identical appearing placebo) in doses between 1200 and 2400 mg daily for 12 weeks, including a 6-week titration phase (95). A significantly larger pain reduction was observed in the gabapentin group, with a mean difference between groups of -0.95 (95% CI: – 1.75 to -0.71; p=0.015). This difference equals an effect size of 0.63 (Cohen's d). A 30% pain reduction was found in 51% compared to 31% in the placebo group (p=0.014, RR was not calculated). No data regarding 50% responder indices were available. Withdrawals due to adverse events were distributed with 16% in the gabapentin group and 10% in the placebo group.

The efficacy of pregabalin on fibromyalgia pain has been investigated in a 2016 Cochrane review, which included five studies with 3283 participants (Table 4-1) (96). A daily dose of 450 mg pregabalin was associated with the best risk-benefit profile. At this dose, 30% pain relief was observed in 43% compared to 29 % in the placebo group, with an RR = 1.5 (95% CI: 1.3 to 1.7). The corresponding proportions for 50% pain relief were 24% versus 14%, with an RR = 1.8 (95% CI: 1.4 to 2.0). Withdrawals due to adverse events occurred in 17% of the pregabalin group versus

9% of the placebo group. The difference in mean pain intensity between groups was not included as an outcome in this review.

Table 4-1. Overview of the evidence for a pain effect of recommended pharmacological treatments in patients with fibromyalgia

Outcome and subgroup	No studies	No patients	Effect size	Quality of evidence
1. Self-reported mean pain intensity				
1.1 Review Amitriptyline	4	275	No data	Very low
1.2 Review Duloxetine	7	2642	-0.26	High
1.3 Review Gabapentin	1	150	-0.63	Very low
1.4 Review Pregabalin	5	3283	No data	High

Outcome and subgroup	No studies	No patients	% in active group	% in placebo group	Risk ratio (95% CI)	Quality of evidence
2. 30% pain improvement						
2.1 Review Amitriptyline	4	275	No data	No data	No data	Very low
2.2 Review Duloxetine	7	2642	47.6	35.9	1.31 (1.2 to 1.4)	High
2.3 Review Gabapentin	1	150	51	31	No data	Very low
2.4 Review Pregabalin	5	3283	43	29	1.5 (1.3 to 1.7)	High
3. 50% pain improvement						
3.1 Review Amitriptyline	4	275	36	11	3.0 (1.7 to 4.9)	Very low
3.2 Review Duloxetine	7	2642	33.9	23.0	1.5 (1.3 to 1. 7)	High
3.3 Review Gabapentin	1	150	No data	No data	No data	Very low
3.4 Review Pregabalin	5	3283	24	14	1.8 (1.4 to 2.0)	High
4. Adverse events withdrawals						
5.1 Review Amitriptyline	4	275	8	9	No data	Very low
5.2 Review Duloxetine	7	2642	15.5	9.5	No data	High
5.3 Review Gabapentin	1	150	16	10	No data	Very low
5.4 Review Pregabalin	5	3283	17	9	No data	High

4.2. Naltrexone - an old drug with new treatment perspectives

Naltrexone is a semi-synthetic non-selective opioid antagonist, patented in the early 60ies, with a high affinity to MOR and less affinity to delta-opioid and kappa-opioid receptors (97). Naltrexone also has binding potential to the opioid growth factor receptor and TLR-2 and TLR-4 (98, 99).

Naltrexone's biochemical structure resembles naloxone, an antidote for the short-term reversal of opioid overdose. Compared to naloxone, naltrexone has a higher oral bioavailability, is converted into an active metabolite (beta-6-naltrexone), and has a substantially longer half-life of about 13 hours (100). Naltrexone is a neutral antagonist that blocks the effects of opioid analgesics but has

no functional effects on opioid receptors. In doses of 50 mg, naltrexone exerts an almost complete blockade of MOR, thereby cancelling the effects of opioid drugs (100). Naltrexone became commercially available to treat opioid abuse in 1984 but lost its patent the following year. Naltrexone also blocks the rewarding effects of other drugs of abuse and was approved for treating alcohol use disorder in 1995 (101).

The off-label use of LDN began soon after its release to the market. A pioneer in this field was Dr Bihari. However, his work has never been published in peer-reviewed journals. In interviews, Dr Bihari described human experiments where he found an increased release of beta-endorphins following treatment with naltrexone that was equally effective in doses of 50 mg, 10 mg, 5 mg and 3 mg (102). Endorphins influence immune function, and Dr Bihari, who worked with patients with AIDS, reported beneficial effects on immune function from off-label treatment with LDN. Dr Bihari used naltrexone in the 3-5 mg range based on the assumption that higher doses might block the endorphins desired effects (103). Subsequently, anecdotal reports began to emerge of the beneficial impact of LDN on cancers, multiple sclerosis, autoimmune disorders, and fibromyalgia (87).

4.3. Proposed mechanism of action of LDN on fibromyalgia

The effects of LDN on fibromyalgia have been suggested to be mediated via 1) a rebound up-regulation of endogenous opioid ligands and receptors or 2) an attenuation of glial reactivity with reduced neuroinflammation (22). The evidence for both hypotheses is very sparse and is briefly reviewed below.

4.3.1. Evidence for the endogenous opioid rebound hypothesis

Animal models have confirmed that chronic administration of naltrexone increases the number of opiate receptors in the brain, and these new receptors are functionally supersensitised (104-106). Furthermore, it has been demonstrated that treatment with naltrexone increases levels of enkephalins in the brains of rats (107). However, in a recent study, treatment with LDN did not result in increased levels of beta-endorphin in the brain or plasma of mice (108). No studies investigating the effect of naltrexone on levels of endogenous opioid receptors or ligands in humans have been published.

4.3.2. Evidence for the anti-inflammatory hypothesis

Activated glial cells are thought to be involved in the pathogenesis of chronic pain and fatigue, and naltrexone has antagonistic properties on TLR-2 and TLR-4 found in glial cells (99). The (+)- and (-)-naltrexone isomeres possess TLR binding activity. However, the (+)-naltrexone isomere, which is inactive at opioid receptors, has been shown to attenuate neuropathic pain in an animal model (109). In a small human study with 8 participants, treatment with LDN 4.5 mg was associated with decreased pain and reduced levels of a range of pro-inflammatory biomarkers (110). However, the cytokines measured were not specific for glial activation.

4.3.3. Considerations about dosing

In the literature, LDN typically refers to doses in the 1-5 mg range, and all clinical trials investigating the efficacy of LDN for Crohn's disease, multiple sclerosis, and fibromyalgia have used test doses between 2 to 4.5 mg (111). This dosing range has not been based on evidence regarding dose and effect but on anecdotal reports. This dosing paradigm is based on the endogenous opioid rebound hypothesis, which assumes that higher doses of naltrexone will block the desired effects (103).

A classic pharmacological dose-effect curve is the S-shaped curve, with a threshold level at which a response first appears, followed by a linear curve where increasing doses result in increasing effects, and finally, a ceiling level where increasing doses do not result in additional effects (112). U-shaped or inverted U-shaped dose-response curves has also been described (112). Naltrexone is primarily known as a neutral antagonist, which cancels the effects of opioids. Initial pharmacodynamic studies have shown increasing antagonistic effects at MOR with increasing doses (97, 100).

The endogenous opioid rebound hypothesis assumes there is a ceiling effect in the low dose range or an inverted U-shaped dose-effect curve. Currently, no evidence exists to support this and no previous studies have investigated the relationship between the dose of naltrexone and its effect on fibromyalgia. Apart from the dose-effect curve, the therapeutic interval of the drug will also depend on the threshold for toxicity. The toxicity threshold for naltrexone is well-examined and lies above the marketed doses of 50 mg (113, 114).

4.4. Previous evidence for the efficacy of LDN in treating fibromyalgia

Before initiating the current LDN trials, we identified two studies with data from a total of 41 participants (published by Younger et al. in 2009 and 2013) which had investigated the efficacy of LDN 4.5 mg compared to a placebo drug for treating fibromyalgia (24, 25). While conducting the current LDN trials, the author of this thesis collaborated in a systematic review on the efficacy of LDN for treating fibromyalgia, which was published in 2023 (Partridge et al.) (26). This systematic search revealed unpublished data by Abou-Raya et al. (a conference abstract from 2013) from an RCT using a parallel design with 74 participants and a test dose of 4.5 mg (115). Most recently, and after the completion of the trials described in this thesis, data from a Danish RCT with a cross-over design and a test dose of 4.5 mg has been published by Bested et al. (116). The evidence from these previous four trials is described below. An overview is provided in Table 4-2.

4.4.1. The Younger et al. 2009 trial

The first trial investigating the potential efficacy of LDN for the treatment of fibromyalgia was a single-blind pilot trial with ten participants (24). All participants received a placebo for two weeks, followed by treatment with LDN 4.5 mg for eight weeks. The participants were not told when or how long to receive active treatment. The primary outcome was a self-reported global measure of fibromyalgia symptom severity using a 101-point visual analogue scale (VAS). Secondary patient-reported outcome measures included daily pain, highest pain, fatigue, sadness, stress, sleep quality, ability to think and remember, gastrointestinal symptoms, and headaches. The results showed that overall fibromyalgia symptoms, daily pain, highest pain, fatigue and stress were significantly more reduced during the LDN condition compared to the placebo condition. No data regarding mean pain reduction or pain responder indices was reported in the manuscript. The results from this trial have been assessed in the Partridge et al. review to have a high risk of bias due to several factors, e.g. a very small sample size, a single-blind design, no random allocation, and no wash-out phase between interventions (26).

4.4.2. The Younger et al. 2013 trial

In 2013, Younger et al. published their second trial investigating the efficacy of LDN for treating fibromyalgia (25). This trial had a double-blind, randomised, placebo-controlled, cross-over design and included 31 women with fibromyalgia. Patients were randomly assigned to four weeks of treatment with a placebo followed by twelve weeks with LDN 4.5 mg or the opposite order with no wash-out period between the two conditions. Two participants dropped out of the treatment, one

because of side effects. A third participant was excluded because of the loss of baseline data. The primary outcome was self-reported daily pain intensity measured using a 101-point VAS scale, and the endpoint was based on average pain for the past three days. Secondary patient-reported outcomes comprised life satisfaction, mood, sleep quality and fatigue. The results showed that pain was averagely reduced by -15.5 (18.0 % ± 10.8%) versus -11.2 (28.8% ± 9.3%) during the placebo and LDN conditions, respectively. This difference was reported to be statistically significant. A standard deviation was not provided for the baseline levels of pain, thus, an effect size could not be estimated from the data. Significant improvements in life satisfaction and mood were observed among the secondary outcomes. A responder was defined as a participant who reported a minimum 30% pain reduction combined with a minimum 30% improvement in sleep quality or fatigue. Based on this definition, nine (32%) of 28 participants were classified as responders. Although the design was double-blind and treatment was randomly allocated, several factors, e.g. lack of a sample size calculation, a small sample size, no wash-out phase between interventions, and exclusion of three participants from the analyses, entailed that the study was assessed to have a high risk of bias in the Partridge et al. review (26).

4.4.3. The Abou-Raya 2013 trial

Data from the third trial investigating the efficacy of LDN for treating fibromyalgia has never been published in a peer-reviewed journal, and data are only available from a conference abstract (115). Efforts from the Partridge group to contact the researchers for more details have been unsuccessful (26). This trial was designed as a parallel RCT where 74 patients (no information about gender) with fibromyalgia were randomly allocated 1:1 to treatment with LDN 4.5 mg or a placebo drug. The primary outcome was a change in self-reported daily pain measured by a 101-point VAS. It was not indicated if the baseline or endpoint measure was average over a defined period. Minimal information was given from the abstract, e.g. no participant flow was provided, drop-outs were not reported, and secondary outcomes were not rigorously defined. Thus, the quality of the methodology can not be sufficiently assessed, and the results were considered to have a high risk of bias by Partridge et al. (26).

4.4.4. The Bested 2023 trial

A fourth study was published in 2023 by Bested et al. with results from an RCT testing the efficacy of LDN for fibromyalgia (116). A cross-over design was applied, and participants of both genders were allocated to three weeks of treatment with LDN 4.5 mg, following three weeks of treatment

with placebo, or the opposite order. Between the two conditions, a wash-out period of two weeks was incorporated. The trial was initially designed as a two-centre trial, and based on information from clinicaltrials.gov, it was intended to include 140 participants. However, patients were only included at one site with 58 participants. A rationale for a sample size of 140 participants is not provided in the manuscript, but a sample size of 51 is argued to be sufficient to estimate a difference between groups equivalent to an effect size of 0.61. Six patients dropped out and were not included in the final analysis. Two primary outcomes were defined a priori as mean changes in 1) FIQR total score and 2) summed scores of pain during rest, personal hygiene measures, and activities of daily living on a 0-30 NRS, also named Summed Pain Intensity Rating (SPIR), measured as the average pain intensity during the past three days. Secondary outcomes included several miscellaneous questionnaires. Pain responder indices were not included as a supportive outcome measure. No between-group differences were observed for this study's primary or secondary outcome. The effect size for the pain outcome was very small, corresponding to a Cohen's d of 0.04.

Table 4-2. Overview of the evidence from four previous LDN/fibromyalgia efficacy trials using a similar test dose of 4.5 mg.

Study characteristics	Younger 2009	Younger 2013	Abou-Raya 2013	Bested 2023
Number of participants randomised/analysed	12/10	31/28	74/74	58/52
Design	Single-blind Cross-over	RCT Cross-over	RCT Parallel	RCT Cross-over
Number of women/men	10/0	28/0	No data	46/6
Treatment in weeks placebo/active	2/8	4/8	24/24	3/3
Wash-out between interventions in cross-over trials in weeks	0	0	N/A	2
Primary outcome	Overall self-reported fibromyalgia symptom severity	Self-reported daily pain Baseline: average 14 days Endpoint: average 3 days	Self-reported daily pain	FIQR total score *and* Summed Pain Intensity Rating (SPIR*) average 3 days

*SPIR = summes three subscores of pain during rest 0-10, personal hygiene measures 0-10, and activities of daily living 0-10 to a 0-30 score

Pain outcomes	Younger 2009	Younger 2013	Abou-Raya 2013	Bested 2023
Change between groups in mean pain	No data	-4.3 (0-100 VAS)	-9.4 (0-100 VAS)	-0.23 (0-30 NRS)
Effect size	No data	No data	-0.69	-0.04
30% pain responders in LD vs placebo groups	No data	33% vs 11%*	No data	No data
50% pain responders in LDN vs placebo groups	No data	No data	No data	No data

*30 % responder criteria in the Younger 2013 trial also required 30% response in fatigue or sleep problems

4.4.5. Core domain set of outcomes for fibromyalgia trials

The core domain set of outcomes for fibromyalgia trials has been defined by the Outcome Measures in Rheumatological Clinical Trials (OMERACT) guidelines (74) and recommends measures of pain, tenderness, fatigue, sleep disturbance, multidimensional function, and patient global in all fibromyalgia trials. The measurement of depression and dyscognition is recommended in some trials, whereas measures of stiffness, anxiety, CSF biomarkers, and functional imaging are optional. Symptom items from the FIQR have commonly been used to measure pain, tenderness, fatigue, sleep disturbance, depression, and dyscognition in fibromyalgia trials. Multidimensional function and patient global have commonly been assessed by the FIQR function domain and FIQR total score, respectively. At least moderate correlations exist between FIQR items/domains and Patient Global Impression of Change (PGI-C) except for FIQR depression (74).

Chapter 5. Current evidence for the efficacy of naltrexone for treating fibromyalgia

Study II aimed to find an optimal naltrexone dose suitable for an efficacy trial. We designed a prospective dose-response study using the up-and-down method for that purpose.

Study III aimed to investigate if a 12-week treatment with LDN was more efficacious than a placebo in reducing pain in patients with fibromyalgia. We designed a randomised, double-blind, placebo-controlled superiority trial to do this.

The study designs, main findings, methodological considerations and implications of the results are briefly reviewed below. Please refer to **Paper II** for further details about **Study II**. For more information on the methodology in **Study III**, please refer to **Paper III**. For a detailed description of the results from **Study III**, please refer to **Paper IV**.

5.1. Methods and materials in study II

The participants for **Study II** were recruited among patients referred to the Pain Centre at Odense University Hospital who had completed a treatment course with insufficient benefits from recommended pharmacological treatments. Enrolment took place between June 2017 and September 2018. Patients eligible were Caucasian women aged 18-60 with an established FM diagnosis at referral. The patients were required to fulfil the ACR1990 and the ACR2010 criteria for fibromyalgia at the time of inclusion. Participants should be able to write and understand Danish and fertile women had to use secure anticonception. Exclusion criteria comprised known allergy towards naltrexone, pregnancy or breastfeeding, disorders of abuse, psychiatric comorbidity, suicide ideation, history of a suicide attempt during the last five years, inflammatory diseases, neurological diseases, and significant localised pain conditions. The use of opioids was prohibited eight weeks before and during the trial.

5.1.1.The up-and-down method

The up-and-down method is a prospective sequential method often used in anaesthesia research to characterise the tolerance distribution of a drug (117). The up-and-down method was originally designed to investigate toxicity in phase I trials, as the method requires fewer participants to be exposed to toxic doses (118). Instead of testing increasing doses in several patients, the individual subject is only exposed to one dose, and the corresponding response is used to determine the test

dose in the following subject. The method allows for estimating a distribution of doses that are precisely sufficient (binary response) to produce an effect among a number of patients. The first patient receives a randomised test dose. In case of a positive response, the dose is lowered for the next participant. In case of no response, the dose is instead increased. The median effective doses for different percentages of the population can be estimated from the dose distribution. Evidence shows that this method can provide a valid estimate of the median effective dose in 50% (ED50) and 95% (ED95) of the population when the dose has shifted direction ten times or when six pairs of up-and-down data are available, requiring about 20-30 participating subjects (117).

Before initiating this study, our clinical practice was to start LDN treatment with a dose of 0.75 mg and to increase the dose with 0.75 mg intervals (based on tolerability) to a maximum dose of 4.5 mg. For **Study II**, we chose a dose range between 0.75 mg and 6 mg, with a dosing interval of 0.75 mg (Figure 5-1). The first participant was randomised to one of 4 doses from the middle of the range (2.25 mg, 3 mg, 3.75 mg or 4.5 mg). The participants were blinded regarding which dose they received.

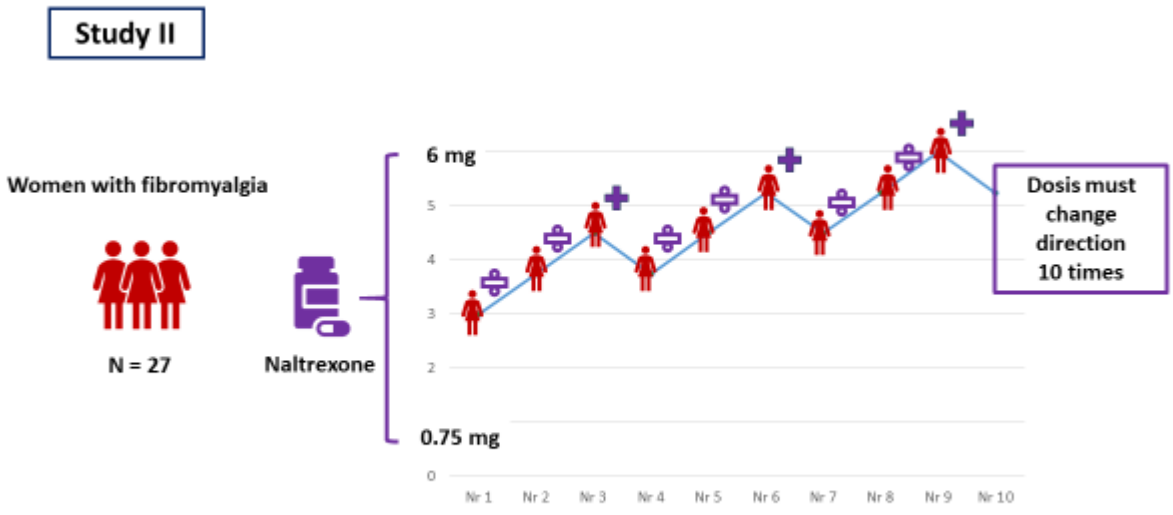


Figure 5-1. Scematic overview of the study flow in study II

5.1.2. Primary outcomes

The up-and-down method requires a binary outcome measure that assesses if the drug "exactly" produces an effect in the subject. For this purpose, we chose two outcomes, requiring only one to be positive. These outcomes were: 1) Patient Global Impression of Improvement (PGI-I) using a 7-point transition scale ranging from 1=very much improved to 7=very much worse, with 4 being no change, and 2) change in average pain during the past three days using a 0-100 visual analogue scale (VAS). The effect was assessed after two weeks of treatment, and a PGI-I score of 1-3 *or* a 30% reduction of pain was considered a positive effect.

5.1.3. Other assessments

Demographic data about age, co-morbidity, and medication were collected. The FIQR symptom domain was used to assess overall symptom severity, and the ten FIQR symptom items were used to explore which symptoms were potentially influenced by LDN. The insomnia severity index (ISI) was used to evaluate changes in sleep quality.

5.1.4. Materials

All data were collected using an electronic data capturing tool (RED-cap). Electronic surveys were sent by e-mail and completed at home by the participants before follow-up visits to reduce bias. Data was transferred to the programme Stata 15 for statistical analysis.

5.1.5. Statistical analyses

Based on the implicit assumption that increasing the dose will increase the effect, isotonic regression analysis was applied to estimate the median effective dose in 50% and 95% of the participants with their 95% confidence limits (117, 118). The Stata 15 IRAX module was used to carry out this analysis. As this method does not allow for imputation of missing data, subjects who were withdrawn or dropped out were replaced.

5.2. Main findings of Study II

A total of 27 participants were included in the study. Two subjects dropped out because of intolerable side effects and were replaced. After the inclusion of 25 evaluable patients, the dose had shifted direction 10 times, and six pairs of up-and-down data were available, enabling the calculation of the ED50 and ED95 based on the dose distribution of the sample. The top of the dosing interval was reached once during the trial, as one participant did not respond to treatment with 6 mg. Consequently, the subsequent patient also received 6 mg.

The median effective dose in the present sample was estimated as follows:
ED50 = 3.9 mg (95% CI 3.4 to 4.4)
ED95 = 5.4 mg (95% CI 4.7 to 6.1)

Among the 25 participants, 11 were assessed as responders to the dose tested. To be assessed as a responder, five subjects fulfilled the pain criteria, ten subjects fulfilled the PGI-I criteria, and four subjects fulfilled both criteria.

For the 11 responders, 30% responder indices were calculated for all ten FIQR items, the FIQR total score, and the ISI score.
After 2-week treatment, the number of responders were:

- 8 of 11 responders for FIQR 'sleep quality'
- 6 of 11 responders for FIQR 'depression'
- 5 of 11 responders for FIQR 'tenderness'
- 4 of 11 responders for FIQR 'pain', 'energy', 'memory' *and* for the ISI score
- 3 of 11 responders for FIQR 'anxiety'
- 2 of 11 responders for FIQR 'stiffness', 'sensory sensitivity' and for the FIQR total score
- 1 of 11 responders for FIQR 'imbalance'

Patients receiving 6 mg reported the least number of side effects per participant (1 side effect reported by 3 subjects). Most side effects per participant were reported in doses below 4.5 mg. No serious adverse events were reported.

5.3. Methodological considerations of Study II

The reason for choosing the up-and-down method to explore the dose distribution of naltrexone was based on our clinical experience that patients with fibromyalgia are susceptible to unwanted treatment effects, and this method allowed us to expose the individual participant to a minimum disadvantage. Using this method, an outcome that assesses if a dose is "precisely" sufficient to give an effect must be defined. Based on our clinical experience as well as evidence from previous trials, we expected that LDN could affect different fibromyalgia symptoms, including pain, sleep, energy, etc. Thus, we used a global transition scale as the primary outcome. The additional pain outcome was chosen, as we planned to include pain as the primary outcome in a subsequent RCT. We found that only five of 11 responders satisfied the pain outcome. However, a less than 30% pain reduction in retrospect could be considered sufficient to be interpreted as a "precisely" positive reaction.

Since we designed the current LDN trials, we have gained experience using LDN in doses up to 9 mg, and this "new" LDN dosing range has also recently been supported by the literature (119). When planning the study, we did not have either evidence or clinical experience with doses higher than 4.5 mg and were reluctant to go higher than 6 mg. No evidence was available to support a ceiling effect at 4.5 mg, and we assumed there would be a linear dose-effect relationship in the dose interval from 0.75 mg to 6 mg. During the trial, we found no problems with tolerability at the top of the dosing interval, and three participants were exposed to 6 mg, hereof one without a treatment effect. Thus, our effective dose estimates might have been higher if we had used a dosing interval of up to 9 mg.

5.4. Implication of findings from Study II

Based on our findings in **Study II**, we concluded that the test dose of 4.5 mg LDN, used in previous trials, was lying in the interval between our estimates of ED50 and ED95. However, if no toxicity is observed at the top of the dosing interval, it could be argued that using the median dose effective in 95% or more of the population could be more optimal. As no problems with tolerability were observed in the current study, we chose the test dose for the subsequent RCT to be 6 mg. However, as efficacy and tolerability might vary due to differences in bioavailability, we also decided to include a titration phase, allowing for slower increments.

By analysing responder indices for other outcomes, we aimed to qualify an appropriate primary outcome for a subsequent RCT. We found that improvement of tenderness was more frequent than pain improvement. This finding could support the hypothesis that LDN primarily influences central sensitisation, with a secondary impact on pain. The symptom most frequently improved was the FIQR 'sleep quality' item. It is well known that pain and sleep affect each other bidirectionally, and sleep improvement could be secondary to the improvement of pain and tenderness. In conclusion, findings from study II did not convince us to deviate from our original aim of investigating the analgesic properties of LDN in the subsequent RCT.

The primary aim of **Study II** was to function as a feasibility study prior to an RCT. We suggest future studies to further explore the dose-effect distribution of naltrexone for treating fibromyalgia, including characterising a possible threshold and ceiling dose.

5.5. Methods and materials in study III

Study III (Fibromyalgia and Naltrexone: the FINAL trial) was designed as a single-centre, randomised, double-blind, placebo-controlled superiority trial with a 1:1 allocation of the participants to treatment with Naltrexone 6 mg once daily or an identically appearing placebo tablet (Figure 5-2). The treatment period was 12 weeks, including a 4-week titration phase. Dosing started at 1.5 mg once daily and was increased by 1.5 mg every week to a maximum dose of 6 mg once daily at week four.

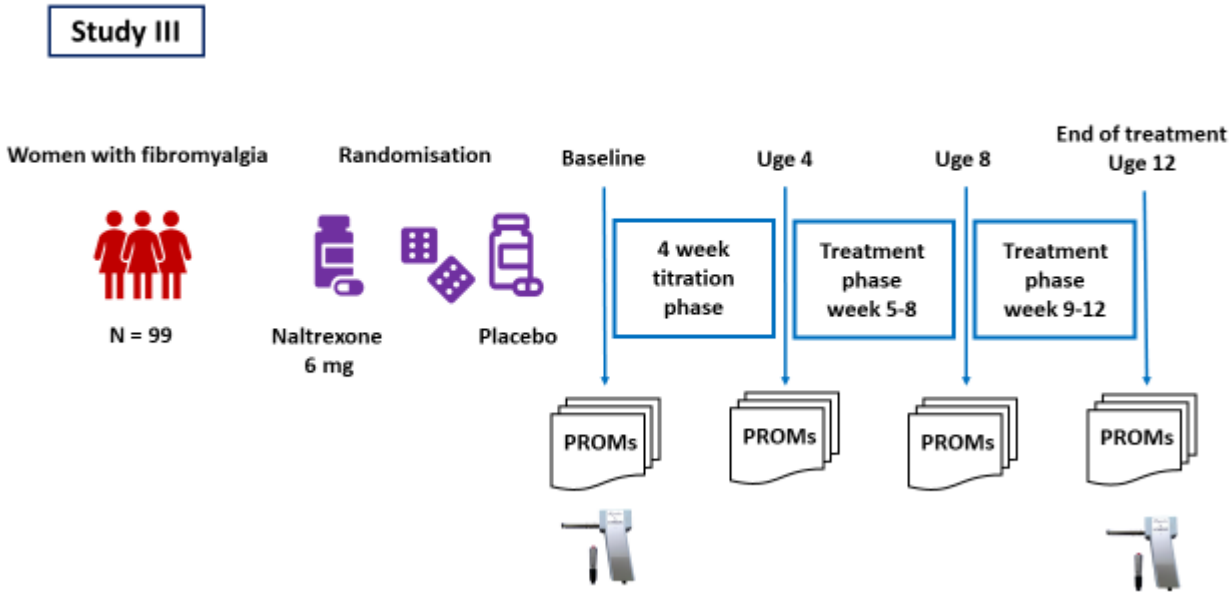


Figure 5-2. Scematic overview of the study flow in study III

The participants for **Study III** were recruited with help from two national patient associations (The Danish Rheumatism Association and The Danish Fibromyalgia and Pain Association) through advertisements in printed and internet-based magazines of these organisations. Patients eligible were women aged 18-64 years with fibromyalgia who fulfilled the ACR1990 criteria and the 2016 diagnostic criteria, who were able to read and write Danish, and who reported pain of at least moderate severity (minimum 4 on a 0-10 NRS, during the last week). Fertile women had to use secure anticonception. Exclusion criteria comprised known allergy towards naltrexone, pregnancy

or breastfeeding, disorders of abuse, known inflammatory rheumatic diseases, known demyelinating diseases, active cancer, psychosis, suicide ideation, known history of suicide attempts, liver dysfunction, and kidney dysfunction. The use of opioids and NSAIDs was prohibited four weeks before and during the trial.

5.4.1. Outcome measures

The primary outcome was the change in pain intensity from baseline to after 12 weeks of treatment. The pain intensity was assessed using the FIQR-pain question, which measures the average pain during the past seven days on a 0-10 NRS.

Secondary outcomes were chosen in accordance with the core set of outcomes for fibromyalgia recommended by OMERACT (74). FIQR single items were used to measure levels of tenderness, fatigue, energy, sleep disturbance, depression, anxiety, memory problems, and stiffness. Physical function was assessed using the FIQR function domain. Patient Global was evaluated using the FIQR total score. The patient's global impression of change (PGI-C) was measured using a 7-point verbal transition scale. Health-related quality of life was assessed using EuroQol-5D (EQ-5D) and EuroQol-VAS (EQ-VAS), a generic instrument that is validated for use in population-based studies and the study of various acute and chronic diseases and is available in a Danish version (120-122).

The pressure pain threshold was assessed using a handheld algometer (Somedic Algometer, Hørby, Sweden). Assessment sites were the right quadriceps muscle, 15 cm from the apex patella and the left trapezius muscle, 10 cm from the acromion (between acromion and C6/7). Each site was assessed three times, and the average of the six values was reported.

Finally, a supportive outcome was included, investigating the number of responders in both treatment groups. The number of responders with a more than 15%, 30%, and 50% improvement of the primary outcome was calculated, with the 30% pain improvement, defined by IMMPACT (the Initiative on Methods, Measurement, and Pain Assessement in Clinical Trials) guidelines as a clinically meaningful improvement, being of primary interest (123).

5.4.2. Harms

Data about harm were collected throughout the trial using active and passive methods. A questionnaire asking for known side effects was administered at all visits based on harm data from **Study II**. Furthermore, participants were encouraged to report any undesired treatment effects throughout the trial.

5.4.3. Other assessments

Tertiary outcomes included measures of pain sensitivity and central pain processing using computerised cuff algometry, including measurements of CPM and TSP. Measures of muscular fatigue comprised the 30s stand chair test (124) and an isometric exhaustion test of the deltoid muscle with electromyographic recordings (125). Finally, a biobank with blood samples was built for later analysis of biomarkers of neuroinflammation and neurotransmitters involved in central pain regulation. These explorative outcomes will be reported in subsequent papers and lie outside the scope of this thesis.

5.4.4. Statistical analyses

We used data from **study II** to calculate an a priori sample size for study **III**. The average pain intensity measured from 0-10 points was estimated to have a mean of 6.7 points and a standard deviation of 1.5 in the target population. Aiming to detect a difference in average pain between groups of 1.0 points with a statistical power of 80% and a significance level of 0.05, a sample size of 74 patients would be required. With the inclusion of about 100 patients in the intention to treat population (approximately 50 in each group), a statistical power of 90% was achieved (126).

The primary analyses were based on the intention-to-treat (ITT) population, comprising all randomised patients. The main analyses were an estimate of the between-group differences for the primary and secondary outcomes after 12 weeks of treatment. For the continuous outcomes, a repeated measures mixed effects model was applied. Estimates were reported as least square means with their 95% CI for each group and the difference between groups, including adjustments for variations in baseline levels. Responder indices were analysed as binary endpoints, comparing the number of responders in the LDN and placebo groups, and were reported as Risk Ratios (126).

Sensitivity analyses were made based on the per-protocol (PP) population, defined as participants with an adherence to the treatment of at least 80%.

5.6. Main findings of Study III

Enrolment took place from January 2021 to December 2022, where 158 patients were screened for eligibility, excluding 59 and allocating 99 patients randomly to treatment with LDN (n=49) or placebo (n=50). There were no dropouts, and the primary and secondary outcomes were assessed for the entire ITT population. The PP population comprised 90 participants (46 in the placebo group

and 44 in the LDN group) with more than 80% adherence to the treatment. The participant flow is visualised in Figure 5-3.

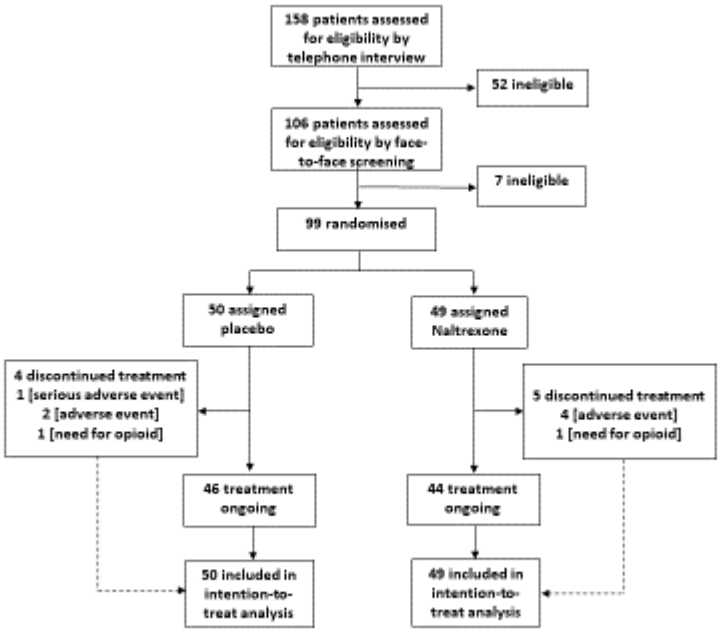


Figure 5-3. Participant flow for the FINAL study

5.6.1. Findings for the primary outcome

The current RCT was not able to confirm the hypothesis that treatment with LDN 6 mg for 12 weeks had a superior effect in reducing pain compared to placebo in women with fibromyalgia.

The trajectories for the primary outcome for the ITT population are visualised in Figure 5-4.

For the **ITT population**, the between-group difference was -0.34 (-0.95 to 0.27; p=0.27) in favour of LDN, corresponding to a Cohen's d of 0.23.

For the **PP population**, the between-group difference was -0.47 (-1.11 to 0.18; p=0.15) in favour of LDN, corresponding to a Cohen's d of 0.31

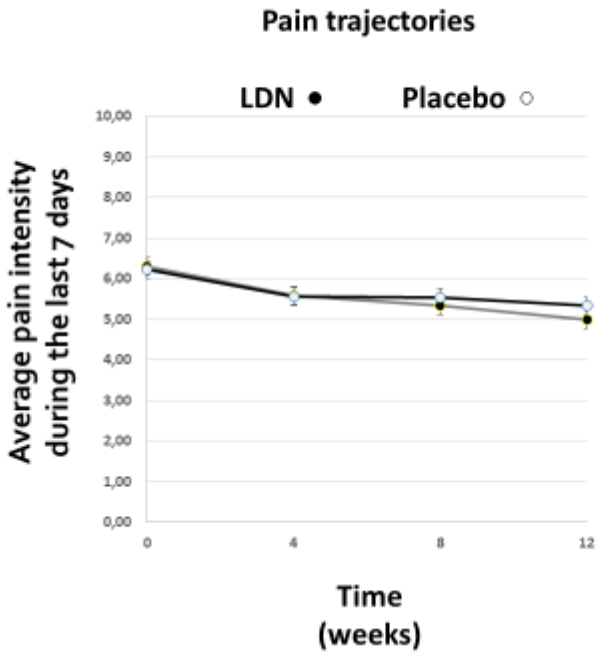


Figure 5-4. Trajectories of the primary outcome over time from baseline to 12 weeks

5.6.2. Findings for the secondary outcomes

For the 13 secondary outcomes, a statistically significant difference between groups was only observed for memory problems in favour of LDN, with a between-group difference of -0.93 (-1.57 to -0.30; p=0.004) in the **ITT population** and -1.01 (-1.69 to -0.34; p=0.004) in the **PP population**.

5.6.3 Responder indicies

A statistical difference between groups for responder indices was not observed. However, for the 30% responders, being of primary interest, the difference between groups approached a significant level for the per-protocol population.

For the **ITT population,** the 30% pain response risk ratio was 1.57 (95% CI 0.88 to 2.79; p=0.12), favouring LDN.

For the **PP population**, the 30% pain response risk ratio was 1.61 (95% CI 0.92 to 2.82; p=0.09), favouring LDN.

5.6.4. Harms

No significant differences were observed between the groups in the overall reporting of adverse events. However, some common adverse events were reported more frequently in the LDN group, including vivid dreams, diarrhoea, dizziness, and hot flashes. Withdrawal from the study medication due to adverse events occurred in 3 (6%) of 50 participants in the placebo group and 4 (8%) of 49 participants in the LDN group. Only one serious adverse event occurred in the placebo group.

5.7. Methodological considerations and interpretations

Study III was designed, conducted, and reported based on the highest standards for clinical trials, adhering to SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) and CONSORT (Consolidated Standards of Reporting Trials) guidelines (127), thus representing the most rigorous LDN/fibromyalgia trial published to date. The current trial was also the first to explore the efficacy of 6 mg LDN on fibromyalgia pain.

The study was powered to detect a difference between groups greater than 1·0 NRS, corresponding to a large effect size. Even though our hypothesis was rejected, several considerations can be made based on the estimates derived from this trial:

1. The 95% confidence interval around the primary estimand excludes a difference greater than 1·0 NRS, equivalent to a large effect size, regarding improving pain.
2. A minimal clinical difference (MID) between groups for the “level of pain” smaller than 1·0 NRS could be clinically relevant. If a distribution-based method was used, it could be suggested that the MID is 0.75 NRS points (1.50/2). The 95% CI around our estimate for the pain outcome does not exclude a difference of 0.75 NRS; thus, there could be a slight chance that our findings are based on a type-2 error.
3. Supposing our estimate of the difference between groups is the actual difference, the study showed a positive treatment effect estimate without a clinically relevant magnitude. A trial would need at least 614 patients (randomised 1:1) to statistically detect a difference between groups corresponding to -0.34 NRS points (using a pooled SD of 1.5).

The study was not powered to detect a difference between groups in responder indices, and we could not demonstrate a significant difference for any responder category. However, when looking at the 95% confidence interval around the estimand for the 30% responder indices, this could be

interpreted as a potentially significant difference, indicating that some patients with fibromyalgia might benefit from LDN treatment.

The only secondary outcome that was significantly improved in the current RCT was memory problems. However, including 16 secondary outcomes (13 key secondary outcomes and three responder indices) increased the chance of a positive finding. Accordingly, the difference between groups regarding improving memory problems was no longer significant when adjusting for multiplicity.

5.8. Implication of finding in Study III

Previous trials investigating the efficacy of 4.5 mg LDN for the treatment of fibromyalgia, showing a potentially large effect size for the reduction of pain, were potentially biased by methodological weaknesses, as described earlier in this thesis. On the contrary, a recent trial also investigating the efficacy of 4.5 mg LDN found a very low effect size regarding their pain outcome. However, as described previously, this trial also had several methodological weaknesses. The current RCT used a very rigorous methodology and thus provides a more robust estimate of the potential efficacy of LDN on fibromyalgia pain than any previous trials.

The current trial rules out a large effect size from treatment with LDN on fibromyalgia pain. However, we found a potentially significant difference between groups regarding 30% pain responders. Supposing the estimates from the current RCT reflect the true treatment effects, an estimated number needed to treat (NNT) about 6-7 can be calculated based on the 30% responder indices and based on the number of withdrawals due to adverse events, numbers needed to harm (NNH) can be estimated to be about 50. Currently, amitriptyline, gabapentin, pregabalin and duloxetine are recommended for the treatment of fibromyalgia. The current thesis provided a review of the evidence for the risks and benefits of these treatments. The beneficial treatment effects found in efficacy trials of amitriptyline and gabapentin are most likely overestimated due to the very low quality of the evidence (90, 94). The evidence for the efficacy of duloxetine and pregabalin has been supported by meta-analyses based on several large RCTs, and the quality of this evidence is high. No general effect has been observed for these two treatments, and the effect size for duloxetine has been estimated to be low (SMD = -0.26). Regarding at least 30% pain relief, the estimated NNT is about 9 for duloxetine and 7 for pregabalin (450 mg) (21, 96). Based on number of withdrawals due to adverse events, the NNH has been estimated to be about 17 for duloxetine

and 11 for pregabalin (21, 96). Thus, LDN could potentially have a better risk-benefit profile than guideline-recommended pharmacological treatments.

Across a range of secondary outcomes in the current RCT, a significant improvement in "memory problems" was observed. However, this could potentially be a false positive finding due to multiplicity. The Younger 2009 trial included the "ability to think and remember" as a secondary outcome, and no significant improvement in this outcome was found. None of the other previous trials have included measures of cognitive dysfunction. Future trials are needed to explore a possible effect of LDN on cognitive function, using validated outcome measures.

In conclusion, this RCT could neither confirm nor rule out that LDN is effective for treating fibromyalgia, and more studies are needed in the future to enable a meta-analysis.

Chapter 6. Conclusions, implications and future perspectives

This thesis aimed to support the identification of fibromyalgia in specialised pain care settings and to fill out the knowledge gap regarding the potential efficacy of LDN for the treatment of fibromyalgia. The results from this thesis have several implications for clinical practice and future research.

In conclusion, the thesis' **Study I** supports that the survey-based 2016 criteria can accurately identify fibromyalgia in clinical practice and for research in specialised pain care settings. Findings from **Study II**, the dose-response study, supported a therapeutic interval of up to 6 mg LDN for the treatment of fibromyalgia. Based on results from **Study III**, LDN was estimated to have a small positive treatment effect on pain without clinically meaningful magnitude. However, a tendency towards a significantly higher number of 30% pain responders in the LDN group compared to the placebo group was observed, with an estimated NNT of 6-7. Both groups had low withdrawals due to adverse events, with an estimated NNH of 50. Thus, compared to guideline-recommended treatments for fibromyalgia, LDN treatment could potentially be associated with a favourable risk-benefit profile.

6.1. Implication of findings

Fibromyalgia has always been a controversial diagnosis, and its identification has relied heavily on the beliefs of patients and physicians. A diagnosis might be influenced by the perception that no effective treatments are available for the disease (128, 129). However, patients with fibromyalgia can be referred to specialised pain care, where pharmacological and non-pharmacological therapies are integrated (130). The FSQ could easily be applied as a screening survey in specialised pain care settings to support the identification of fibromyalgia. A fibromyalgia diagnosis will direct the pharmacological treatment strategy towards central-acting drugs instead of peripherally-acting medications (12). Furthermore, evidence shows that diagnostic certainty attenuates fearful beliefs about pain and supports better self-management (131).

During the past decade, the prescription of LDN for fibromyalgia has increased rapidly, gaining new hope for an effective and safe treatment for fibromyalgia. However, the potential benefits of this treatment were previously only supported by evidence from low-quality trials (26). Results from **Study III** have a low risk of bias due to high robustness regarding methodology, and thus, provide the most valid estimates of the risks and benefits associated with LDN treatment to date.

Even though the evidence from **Study III** impeeds the hope that LDN could be a new, generally effective treatment for fibromyalgia, the findings indicate that the risk-benefit profile could be better for LDN than for guideline-recommended treatments. Thus, findings from this thesis does not support advice against using LDN in specialised pain care settings. However, more trials are still needed before firm recommendations can be made regarding the future use of LDN.

6.2. Future directions

A significant challenge in the treatment of fibromyalgia is that existing pharmacological treatments have no general effect on pain. This might be explained by the fact that many different factors can contribute to the perpetuation of pain across the same phenotype, and better management relies on new treatments targeting underlying mechanisms. Thus, greater insights into fibromyalgia neurobiology are warranted to support future targeted pharmacological therapies. If neurobiological aberrations specific to fibromyalgia are found, this could also lead to the discovery of objective biomarkers to validate a diagnosis of fibromyalgia.

More studies are also needed in the future to comfirm the estimated effects of LDN for the treatment of fibromyalgia found in the current thesis:

- Although the current thesis could not rule out a medium effect size from treatment with LDN, it is most likely that only some patients with fibromyalgia will benefit from LDN treatment. Thus, future efficacy trials should preferably be powered to detect a clinically relevant difference in responder indices instead of group differences.
- Future studies are warranted to explore factors that could predict a positive response to support a personalised treatment approach.
- Future dose-response studies are also highly warranted to explore a therapeutic interval of LDN, including threshold and ceiling levels.
- Future studies might explore if doses higher than 6 mg are more beneficial. Previous trials have only investigated the efficacy of 4.5 mg, and the current RCT was the first to use a test dose of 6 mg. The use of doses up to 9 mg has been supported in the literature (119), and anecdotal reports of even higher doses exist on internet pages.
- Based on the clinical experience with LDN, an improvement of dyscognition and executive dysfunction (fibro-fog) is often reported. Future trials are needed to investigate the potential efficacy of LDN on dyscognition associated with fibromyalgia.

- Evidence for a treatment effect of LDN in men with fibromyalgia is lacking. It can be difficult to recruit men with fibromyalgia for clinical trials, as men are often not appropriately diagnosed in clinical settings. The 2016 criteria could be applied as a tool for the identification of men with fibromyalgia in future research.
- More knowledge is needed regarding the putative mechanisms of action of LDN, as it could potentially guide the development of new and more effective pharmacological treatments.

In the future, and lying outside the scope of this thesis, we will perform responder analyses based on data from the RCT, to explore the hypothesis that an inflammatory subtype of fibromyalgia might benefit from treatment with LDN. Furthermore, possible mechanisms of action of LDN will be examined based on exploratory outcomes from the RCT, including experimental measures pain processing and changes in inflammatory and glial biomarkers.

References

1. Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). Pain. 2019;160(1):19-27.
2. Nicholas M, Vlaeyen JWS, Rief W, Barke A, Aziz Q, Benoliel R, et al. The IASP classification of chronic pain for ICD-11: chronic primary pain. Pain. 2019;160(1):28-37.
3. Mezhov V, Guymer E, Littlejohn G. Central sensitivity and fibromyalgia. Intern Med J. 2021;51(12):1990-8.
4. Freynhagen R, Parada HA, Calderon-Ospina CA, Chen J, Rakhmawati Emril D, Fernandez-Villacorta FJ, et al. Current understanding of the mixed pain concept: a brief narrative review. Curr Med Res Opin. 2019;35(6):1011-8.
5. Fitzcharles MA, Cohen SP, Clauw DJ, Littlejohn G, Usui C, Hauser W. Nociplastic pain: towards an understanding of prevalent pain conditions. Lancet. 2021;397(10289):2098-110.
6. Bidari A, Ghavidel-Parsa B. Nociplastic pain concept, a mechanistic basis for pragmatic approach to fibromyalgia. Clin Rheumatol. 2022;41(10):2939-47.
7. Amris K, Waehrens EE, Jespersen A, Bliddal H, Danneskiold-Samsoe B. Observation-based assessment of functional ability in patients with chronic widespread pain: a cross-sectional study. Pain. 2011;152(11):2470-6.
8. Heidari F, Afshari M, Moosazadeh M. Prevalence of fibromyalgia in general population and patients, a systematic review and meta-analysis. Rheumatol Int. 2017;37(9):1527-39.
9. Jensen HI, Plesner K, Kvorning N, Krogh BL, Kimper-Karl A. Associations between demographics and health-related quality of life for chronic non-malignant pain patients treated at a multidisciplinary pain centre: a cohort study. Int J Qual Health Care. 2016;28(1):86-91.
10. Woolf CJ, Max MB. Mechanism-based pain diagnosis: issues for analgesic drug development. Anesthesiology. 2001;95(1):241-9.
11. Clauw DJ. Fibromyalgia: a clinical review. JAMA. 2014;311(15):1547-55.
12. Tzadok R, Ablin JN. Current and Emerging Pharmacotherapy for Fibromyalgia. Pain Res Manag. 2020;2020:6541798.
13. Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, et al. The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multicenter Criteria Committee. Arthritis Rheum. 1990;33(2):160-72.
14. Fitzcharles MA, Boulos P. Inaccuracy in the diagnosis of fibromyalgia syndrome: analysis of referrals. Rheumatology. 2003;42(2):263-7.
15. Buskila D, Neumann L, Sibirski D, Shvartzman P. Awareness of diagnostic and clinical features of fibromyalgia among family physicians. Fam Pract. 1997;14(3):238-41.
16. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Katz RS, Mease P, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken). 2010;62(5):600-10.
17. Egloff N, von Kanel R, Muller V, Egle UT, Kokinogenis G, Lederbogen S, et al. Implications of proposed fibromyalgia criteria across other functional pain syndromes. Scand J Rheumatol. 2015;44(5):416-24.

18. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Hauser W, Katz RL, et al. 2016 Revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum. 2016;46(3):319-29.
19. Ablin JN, Wolfe F. A Comparative Evaluation of the 2011 and 2016 Criteria for Fibromyalgia. J Rheumatol. 2017;44(8):1271-6.
20. Tzellos TG, Toulis KA, Goulis DG, Papazisis G, Zampeli VA, Vakfari A, et al. Gabapentin and pregabalin in the treatment of fibromyalgia: a systematic review and a meta-analysis. J Clin Pharm Ther. 2010;35(6):639-56.
21. 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(1):71-82.
22. Younger J, Parkitny L, McLain D. The use of low-dose naltrexone (LDN) as a novel anti-inflammatory treatment for chronic pain. Clin Rheumatol. 2014;33(4):451-9.
23. Raknes G, Smabrekke L. A sudden and unprecedented increase in low dose naltrexone (LDN) prescribing in Norway. Patient and prescriber characteristics, and dispense patterns. A drug utilization cohort study. Pharmacoepidemiol Drug Saf. 2017;26(2):136-42.
24. Younger J, Mackey S. Fibromyalgia symptoms are reduced by low-dose naltrexone: a pilot study. Pain Med. 2009;10(4):663-72.
25. 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(2):529-38.
26. Partridge S, Quadt L, Bolton M, Eccles J, Thompson C, Colasanti A, 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(5):e15638.
27. Inanici F, Yunus MB. History of fibromyalgia: past to present. Curr Pain Headache Rep. 2004;8(5):369-78.
28. Csillag C. Fibromyalgia - the Copenhagen Declaration. Lancet. 1992;340(8820):663-4.
29. Siracusa R, Paola RD, Cuzzocrea S, Impellizzeri D. Fibromyalgia: Pathogenesis, Mechanisms, Diagnosis and Treatment Options Update. Int J Mol Sci. 2021;22(8).
30. Marques AP, Santo A, Berssaneti AA, Matsutani LA, Yuan SLK. Prevalence of fibromyalgia: literature review update. Rev Bras Reumatol Engl Ed. 2017;57(4):356-63.
31. Wolfe F, Walitt B, Perrot S, Rasker JJ, Hauser W. Fibromyalgia diagnosis and biased assessment: Sex, prevalence and bias. PLoS One. 2018;13(9):e0203755.
32. Yunus MB. The role of gender in fibromyalgia syndrome. Curr Rheumatol Rep. 2001;3(2):128-34.
33. Bennett RM, Friend R. Secondary Fibromyalgia. J Rheumatol. 2019;46(2):127-9.
34. Yunus MB. The prevalence of fibromyalgia in other chronic pain conditions. Pain Res Treat. 2012;2012:584573.
35. Wolfe F, Ablin J, Guymer EK, Littlejohn GO, Rasker JJ. The Relation of Physical Comorbidity and Multimorbidity to Fibromyalgia, Widespread Pain, and Fibromyalgia-related Variables. J Rheumatol. 2020;47(4):624-31.
36. Sluka KA, Clauw DJ. Neurobiology of fibromyalgia and chronic widespread pain. Neuroscience. 2016;338:114-29.

37. Yunus MB, Kalyan-Raman UP, Masi AT, Aldag JC. Electron microscopic studies of muscle biopsy in primary fibromyalgia syndrome: a controlled and blinded study. J Rheumatol. 1989;16(1):97-101.
38. Simms RW, Roy SH, Hrovat M, Anderson JJ, Skrinar G, LePoole SR, et al. Lack of association between fibromyalgia syndrome and abnormalities in muscle energy metabolism. Arthritis Rheum. 1994;37(6):794-800.
39. Simms RW. Is there muscle pathology in fibromyalgia syndrome? Rheum Dis Clin North Am. 1996;22(2):245-66.
40. Grayston R, Czanner G, Elhadd K, Goebel A, Frank B, Uceyler N, et al. A systematic review and meta-analysis of the prevalence of small fiber pathology in fibromyalgia: Implications for a new paradigm in fibromyalgia etiopathogenesis. Semin Arthritis Rheum. 2019;48(5):933-40.
41. Serra J, Collado A, Sola R, Antonelli F, Torres X, Salgueiro M, et al. Hyperexcitable C nociceptors in fibromyalgia. Ann Neurol. 2014;75(2):196-208.
42. O'Brien AT, Deitos A, Trinanes Pego Y, Fregni F, Carrillo-de-la-Pena MT. Defective Endogenous Pain Modulation in Fibromyalgia: A Meta-Analysis of Temporal Summation and Conditioned Pain Modulation Paradigms. J Pain. 2018;19(8):819-36.
43. 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(1):68-75.
44. Vanneste S, Ost J, Van Havenbergh T, De Ridder D. Resting state electrical brain activity and connectivity in fibromyalgia. PLoS One. 2017;12(6):e0178516.
45. Andersen JV, Schousboe A. Milestone Review: Metabolic dynamics of glutamate and GABA mediated neurotransmission - The essential roles of astrocytes. J Neurochem. 2023;166(2):109-37.
46. Pyke TL, Osmotherly PG, Baines S. Measuring Glutamate Levels in the Brains of Fibromyalgia Patients and a Potential Role for Glutamate in the Pathophysiology of Fibromyalgia Symptoms: A Systematic Review. Clin J Pain. 2017;33(10):944-54.
47. Harris RE. Elevated excitatory neurotransmitter levels in the fibromyalgia brain. Arthritis Res Ther. 2010;12(5):141.
48. Foerster BR, Petrou M, Edden RA, Sundgren PC, Schmidt-Wilcke T, Lowe SE, et al. Reduced insular gamma-aminobutyric acid in fibromyalgia. Arthritis Rheum. 2012;64(2):579-83.
49. Harris RE, Clauw DJ, Scott DJ, McLean SA, Gracely RH, Zubieta JK. Decreased central mu-opioid receptor availability in fibromyalgia. Journal of Neuroscience. 2007;27(37):10000-6.
50. Baraniuk JN, Whalen G, Cunningham J, Clauw DJ. Cerebrospinal fluid levels of opioid peptides in fibromyalgia and chronic low back pain. BMC Musculoskeletal Disorders. 2004;5:48.
51. Donnelly CR, Andriessen AS, Chen G, Wang K, Jiang C, Maixner W, et al. Central Nervous System Targets: Glial Cell Mechanisms in Chronic Pain. Neurotherapeutics. 2020;17(3):846-60.
52. Kumbhare D, Hassan S, Diep D, Duarte FCK, Hung J, Damodara S, et al. Potential role of blood biomarkers in patients with fibromyalgia: a systematic review with meta-analysis. Pain. 2022;163(7):1232-53.

53. Albrecht DS, Forsberg A, Sandstrom A, Bergan C, Kadetoff D, Protsenko E, et al. Brain glial activation in fibromyalgia - A multi-site positron emission tomography investigation. Brain Behav Immun. 2019;75:72-83.
54. Aggarwal R, Ringold S, Khanna D, Neogi T, Johnson SR, Miller A, et al. Distinctions between diagnostic and classification criteria? Arthritis Care Res (Hoboken). 2015;67(7):891-7.
55. Smythe HA, Moldofsky H. Two contributions to understanding of the "fibrositis" syndrome. Bull Rheum Dis. 1977;28(1):928-31.
56. Bennett RM. Fibrositis: misnomer for a common rheumatic disorder. West J Med. 1981;134(5):405-13.
57. Yunus M, Masi AT, Calabro JJ, Miller KA, Feigenbaum SL. Primary fibromyalgia (fibrositis): clinical study of 50 patients with matched normal controls. Semin Arthritis Rheum. 1981;11(1):151-71.
58. Wolfe F, Walitt B. Culture, science and the changing nature of fibromyalgia. Nat Rev Rheumatol. 2013;9(12):751-5.
59. Butler S, Landmark T, Glette M, Borchgrevink P, Woodhouse A. Chronic widespread pain-the need for a standard definition. Pain. 2016;157(3):541-3.
60. Cohen ML, Quintner JL. Fibromyalgia syndrome, a problem of tautology. Lancet. 1993;342(8876):906-9.
61. Crofford LJ, Clauw DJ. Fibromyalgia: Where are we a decade after the American College of Rheumatology classification criteria were developed? Arthritis Rheum-Us. 2002;46(5):1136-8.
62. Wolfe F. Stop using the American College of Rheumatology criteria in the clinic. Journal of Rheumatology. 2003;30(8):1671-2.
63. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Hauser W, Katz RS, et al. Fibromyalgia criteria and severity scales for clinical and epidemiological studies: a modification of the ACR Preliminary Diagnostic Criteria for Fibromyalgia. J Rheumatol. 2011;38(6):1113-22.
64. Hauser W, Jung E, Erbsloh-Moller B, Gesmann M, Kuhn-Becker H, Petermann F, et al. Validation of the Fibromyalgia Survey Questionnaire within a cross-sectional survey. PLoS One. 2012;7(5):e37504.
65. Kim SM, Lee SH, Kim HR. Applying the ACR Preliminary Diagnostic Criteria in the Diagnosis and Assessment of Fibromyalgia. Korean J Pain. 2012;25(3):173-82.
66. Wolfe F, Egloff N, Hauser W. Widespread Pain and Low Widespread Pain Index Scores among Fibromyalgia-positive Cases Assessed with the 2010/2011 Fibromyalgia Criteria. J Rheumatol. 2016;43(9):1743-8.
67. Kang JH, An M, Choi SE, Xu H, Park DJ, Lee JK, et al. Performance of the revised 2016 fibromyalgia diagnostic criteria in Korean patients with fibromyalgia. Int J Rheum Dis. 2019;22(9):1734-40.
68. Fors EA, Wensaas KA, Eide H, Jaatun EA, Clauw DJ, Wolfe F, et al. Fibromyalgia 2016 criteria and assessments: comprehensive validation in a Norwegian population. Scand J Pain. 2020;20(4):663-72.
69. Salaffi F, Di Carlo M, Farah S, Atzeni F, Buskila D, Ablin JN, et al. Diagnosis of fibromyalgia: comparison of the 2011/2016 ACR and AAPT criteria and validation of the modified Fibromyalgia Assessment Status. Rheumatology (Oxford). 2020;59(10):3042-9.

70. Ahmed S, Aggarwal A, Lawrence A. Performance of the American College of Rheumatology 2016 criteria for fibromyalgia in a referral care setting. Rheumatol Int. 2019;39(8):1397-403.
71. Wolfe F, Walitt BT, Rasker JJ, Katz RS, Hauser W. The Use of Polysymptomatic Distress Categories in the Evaluation of Fibromyalgia (FM) and FM Severity. J Rheumatol. 2015;42(8):1494-501.
72. Bennett R. The Fibromyalgia Impact Questionnaire (FIQ): a review of its development, current version, operating characteristics and uses. Clin Exp Rheumatol. 2005;23(5 Suppl 39):S154-62.
73. 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(4):R120.
74. Mease PJ, Clauw DJ, Christensen R, Crofford LJ, Gendreau RM, Martin SA, et al. Toward development of a fibromyalgia responder index and disease activity score: OMERACT module update. J Rheumatol. 2011;38(7):1487-95.
75. Friend R, Bennett RM. Distinguishing fibromyalgia from rheumatoid arthritis and systemic lupus in clinical questionnaires: an analysis of the revised Fibromyalgia Impact Questionnaire (FIQR) and its variant, the Symptom Impact Questionnaire (SIQR), along with pain locations. Arthritis Res Ther. 2011;13(2):R58.
76. Friend R, Bennett RM, Aebischer JH, St John AW, Jones KD. Validating the revised Symptom Impact Questionnaire with a proposed fibromyalgia phenotype using experimentally-induced pain and patient self-reports. Clin Exp Rheumatol. 2021;39 Suppl 130(3):137-43.
77. Korevaar DA, Cohen JF, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, et al. Updating standards for reporting diagnostic accuracy: the development of STARD 2015. Res Integr Peer Rev. 2016;1:7.
78. Amris K, Waehrens EE, Jespersen A, Stockmarr A, Bennett R, Bliddal H, et al. The Relationship between Mechanical Hyperalgesia Assessed by Manual Tender Point Examination and Disease Severity in Patients with Chronic Widespread Pain: A Cross-Sectional Study. Int J Rheumatol. 2014;2014:417596.
79. Bujang MA, Adnan TH. Requirements for Minimum Sample Size for Sensitivity and Specificity Analysis. J Clin Diagn Res. 2016;10(10):YE01-YE6.
80. Macfarlane GJ, Kronisch C, Dean LE, Atzeni F, Hauser W, Fluss E, et al. EULAR revised recommendations for the management of fibromyalgia. Ann Rheum Dis. 2017;76(2):318-28.
81. Choy E, Marshall D, Gabriel ZL, Mitchell SA, Gylee E, Dakin HA. A systematic review and mixed treatment comparison of the efficacy of pharmacological treatments for fibromyalgia. Semin Arthritis Rheum. 2011;41(3):335-45 e6.
82. Ernst E. Chiropractic treatment for fibromyalgia: a systematic review. Clin Rheumatol. 2009;28(10):1175-8.
83. Li YH, Wang FY, Feng CQ, Yang XF, Sun YH. Massage therapy for fibromyalgia: a systematic review and meta-analysis of randomized controlled trials. PLoS One. 2014;9(2):e89304.
84. Finnerup NB, Kuner R, Jensen TS. Neuropathic Pain: From Mechanisms to Treatment. Physiol Rev. 2021;101(1):259-301.
85. Roskell NS, Beard SM, Zhao Y, Le TK. A meta-analysis of pain response in the treatment of fibromyalgia. Pain Pract. 2011;11(6):516-27.

86. Miotto K, Cho AK, Khalil MA, Blanco K, Sasaki JD, Rawson R. Trends in Tramadol: Pharmacology, Metabolism, and Misuse. Anesth Analg. 2017;124(1):44-51.
87. Brown N, Panksepp J. Low-dose naltrexone for disease prevention and quality of life. Med Hypotheses. 2009;72(3):333-7.
88. Sundhedsstyrelsen. National klinisk retningslinje for udredning og behandling samt rehabilitering af patienter med generalisede smerter i bevægeapparatet sst.dk2018 [
89. Rico-Villademoros F, Slim M, Calandre EP. Amitriptyline for the treatment of fibromyalgia: a comprehensive review. Expert Rev Neurother. 2015;15(10):1123-50.
90. Moore RA, Derry S, Aldington D, Cole P, Wiffen PJ. Amitriptyline for fibromyalgia in adults. Cochrane Database Syst Rev. 2019;5(7):CD011824.
91. Bellingham GA, Peng PW. Duloxetine: a review of its pharmacology and use in chronic pain management. Reg Anesth Pain Med. 2010;35(3):294-303.
92. Fishbain D. Evidence-based data on pain relief with antidepressants. Ann Med. 2000;32(5):305-16.
93. Alles SRA, Cain SM, Snutch TP. Pregabalin as a Pain Therapeutic: Beyond Calcium Channels. Front Cell Neurosci. 2020;14:83.
94. Cooper TE, Derry S, Wiffen PJ, Moore RA. Gabapentin for fibromyalgia pain in adults. Cochrane Database Syst Rev. 2017;1(1):CD012188.
95. Arnold LM, Goldenberg DL, Stanford SB, Lalonde JK, Sandhu HS, Keck PE, Jr., et al. Gabapentin in the treatment of fibromyalgia: a randomized, double-blind, placebo-controlled, multicenter trial. Arthritis Rheum. 2007;56(4):1336-44.
96. Derry S, Cording M, Wiffen PJ, Law S, Phillips T, Moore RA. Pregabalin for pain in fibromyalgia in adults. Cochrane Database Syst Rev. 2016;9(9):CD011790.
97. Crabtree BL. Review of naltrexone, a long-acting opiate antagonist. Clin Pharm. 1984;3(3):273-80.
98. McLaughlin PJ, McHugh DP, Magister MJ, Zagon IS. Endogenous opioid inhibition of proliferation of T and B cell subpopulations in response to immunization for experimental autoimmune encephalomyelitis. BMC Immunol. 2015;16:24.
99. Kwilasz AJ, Todd LS, Duran-Malle JC, Schrama AEW, Mitten EH, Larson TA, et al. Experimental autoimmune encephalopathy (EAE)-induced hippocampal neuroinflammation and memory deficits are prevented with the non-opioid TLR2/TLR4 antagonist (+)-naltrexone. Behav Brain Res. 2021;396:112896.
100. Verebey K. The clinical pharmacology of naltrexone: pharmacology and pharmacodynamics. NIDA Res Monogr. 1981;28:147-58.
101. Ray LA, Chin PF, Miotto K. Naltrexone for the treatment of alcoholism: clinical findings, mechanisms of action, and pharmacogenetics. CNS Neurol Disord Drug Targets. 2010;9(1):13-22.
102. Bihari B. Bernard Bihari, MD: low-dose naltrexone for normalizing immune system function. Altern Ther Health Med. 2013;19(2):56-65.
103. Bihari B. Efficacy of low dose naltrexone as an immune stabilizing agent for the treatment of HIV/AIDS. AIDS Patient Care. 1995;9(1):3.
104. Zukin RS, Sugarman JR, Fitz-Syage ML, Gardner EL, Zukin SR, Gintzler AR. Naltrexone-induced opiate receptor supersensitivity. Brain Res. 1982;245(2):285-92.
105. Tempel A, Gardner EL, Zukin RS. Neurochemical and functional correlates of naltrexone-induced opiate receptor up-regulation. J Pharmacol Exp Ther. 1985;232(2):439-44.

106. Sirohi S, Kumar P, Yoburn BC. Mu-opioid receptor up-regulation and functional supersensitivity are independent of antagonist efficacy. J Pharmacol Exp Ther. 2007;323(2):701-7.
107. Tempel A, Kessler JA, Zukin RS. Chronic naltrexone treatment increases expression of preproenkephalin and preprotachykinin mRNA in discrete brain regions. J Neurosci. 1990;10(3):741-7.
108. 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 beta-Endorphin. eNeuro. 2021;8(3).
109. Hutchinson MR, Zhang Y, Brown K, Coats BD, Shridhar M, Sholar PW, et al. Non-stereoselective reversal of neuropathic pain by naloxone and naltrexone: involvement of toll-like receptor 4 (TLR4). Eur J Neurosci. 2008;28(1):20-9.
110. Parkitny L, Younger J. Reduced Pro-Inflammatory Cytokines after Eight Weeks of Low-Dose Naltrexone for Fibromyalgia. Biomedicines. 2017;5(2).
111. Toljan K, Vrooman B. Low-Dose Naltrexone (LDN)-Review of Therapeutic Utilization. Med Sci (Basel). 2018;6(4).
112. Yang L, Wang J, Cheke RA, Tang S. A Universal Delayed Difference Model Fitting Dose-response Curves. Dose Response. 2021;19(4):15593258211062785.
113. Yoon G, Kim SW, Thuras P, Westermeyer J. Safety, tolerability, and feasibility of high-dose naltrexone in alcohol dependence: an open-label study. Hum Psychopharmacol. 2011;26(2):125-32.
114. Verebey KG, Mule SJ. Naltrexone (Trexan): a review of hepatotoxicity issues. NIDA Res Monogr. 1986;67:73-81.
115. Abou-Raya S, Abou-Raya A, Khadrawi T. Efficacy of naltrexone in the treatment of fibromyalgia: randomized controlled trial. Annals of the Rheumatic Diseases 2013.
116. Bested KJ, L; Andresen, T; Tarp, G; Skovbjerg, L; Johansen, TSD; Schmedes, AV; Storgaard, I; Madsen, J; Werner, MU; Bendiksen, A. Low-dose naltrexone for treatment of pain in patients with fibromyalgia: a randomized, double-blind, placebo-controlled, crossover study. PAIN reports. 2023;8(4):e1088.
117. Pace NL, Stylianou MP. Advances in and limitations of up-and-down methodology: a precis of clinical use, study design, and dose estimation in anesthesia research. Anesthesiology. 2007;107(1):144-52.
118. Stylianou M, Flournoy N. Dose finding using the biased coin up-and-down design and isotonic regression. Biometrics. 2002;58(1):171-7.
119. 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(4).
120. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. Applied Health Economics and Health Policy. 2017;15(2):127-37.
121. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.
122. Stolk E, Ludwig K, Rand K, van Hout B, Ramos-Goni JM. Overview, Update, and Lessons Learned From the International EQ-5D-5L Valuation Work: Version 2 of the EQ-5D-5L Valuation Protocol. Value Health. 2019;22(1):23-30.

123. Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain. 2008;9(2):105-21.
124. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport. 1999;70(2):113-9.
125. Bandak E, Amris K, Bliddal H, Danneskiold-Samsoe B, Henriksen M. Muscle fatigue in fibromyalgia is in the brain, not in the muscles: a case-control study of perceived versus objective muscle fatigue. Ann Rheum Dis. 2013;72(6):963-6.
126. Bruun KD, Amris K, Vaegter HB, Blichfeldt-Eckhardt MR, Holsgaard-Larsen A, Christensen R, et al. Low-dose naltrexone for the treatment of fibromyalgia: protocol for a double-blind, randomized, placebo-controlled trial. Trials. 2021;22(1):804.
127. Hopewell S, Boutron I, Chan AW, Collins GS, de Beyer JA, Hrobjartsson A, et al. An update to SPIRIT and CONSORT reporting guidelines to enhance transparency in randomized trials. Nat Med. 2022;28(9):1740-3.
128. Perrot S, Choy E, Petersel D, Ginovker A, Kramer E. Survey of physician experiences and perceptions about the diagnosis and treatment of fibromyalgia. BMC Health Serv Res. 2012;12:356.
129. Hayes SM, Myhal GC, Thornton JF, Camerlain M, Jamison C, Cytryn KN, et al. Fibromyalgia and the therapeutic relationship: where uncertainty meets attitude. Pain Res Manag. 2010;15(6):385-91.
130. Vaegter HB, Christoffersen LO, Enggaard TP, Holdggard DEM, Lefevre TN, Eltved R, et al. Socio-Demographics, Pain Characteristics, Quality of Life and Treatment Values Before and After Specialized Interdisciplinary Pain Treatment: Results from the Danish Clinical Pain Registry (PainData). J Pain Res. 2021;14:1215-30.
131. Doebl S, Macfarlane GJ, Hollick RJ. 'No one wants to look after the fibro patient'. Understanding models, and patient perspectives, of care for fibromyalgia: reviews of current evidence. Pain. 2020.

Appendix A

REVIDERET FIBROMYALGIA IMPACT QUESTIONNAIRE (FIQR)

Efternavn: **Fornavn:** **Alder:**

Varighed af fibromyalgi symptomer (år)**:** Tid siden fibromyalgi blev diagnosticeret første gang (år):

Instruktion: For hvert af de følgende 9 spørgsmål, sæt kryds i det felt, som bedst beskriver, i hvor høj grad din fibromyalgi gjorde det vanskeligt at udføre hver af de følgende aktiviteter i løbet af de seneste 7 dage. Hvis du ikke har udført en af de angivne aktiviteter de sidste 7 dage, angiv da hvor vanskeligt det var sidste gang, du udførte aktiviteten. Hvis du ikke kan udføre aktiviteten, sæt kryds i det sidste felt.

Børste eller rede dit hår	**Ikke vanskeligt** ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ **Meget vanskeligt**
20 minutters uafbrudt gang	**Ikke vanskeligt** ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ **Meget vanskeligt**
Tilberede et måltid mad	**Ikke vanskeligt** ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ **Meget vanskeligt**
Støvsuge, vaske eller feje gulv	**Ikke vanskeligt** ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ **Meget vanskeligt**
Løfte og bære en fyldt indkøbspose	**Ikke vanskeligt** ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ **Meget vanskeligt**
Tage trapperne en etage op	**Ikke vanskeligt** ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ **Meget vanskeligt**
Skifte sengetøj	**Ikke vanskeligt** ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ **Meget vanskeligt**
Sidde i en stol I 45 minutter	**Ikke vanskeligt** ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ **Meget vanskeligt**
Gå på indkøb efter dagligvarer	**Ikke vanskeligt** ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ **Meget vanskeligt**

Sub-total ***(kun til internt brug)*** ☐

Instruktion: For hvert af de følgende 2 spørgsmål sæt kryds i det ene felt, der bedst beskriver din fibromyalgis samlede indflydelse i løbet af de sidste 7 dage.

Fibromyalgien forhindrede mig i at opfylde mål i løbet af ugen	**Aldrig** ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ **Hele tiden**
Jeg var fuldstændig overvældet af mine fibromyalgi-symptomer	**Aldrig** ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ **Hele tiden**

Sub-total ***(kun til internt brug)*** ☐

Appendix A

Instruktion: For hvert af de følgende 10 spørgsmål sæt kryds i det ene felt, der bedst beskriver intensiteten af de følgende almindelige symptomer i løbet af de sidste 7 dage.

Vurder dit niveau af smerte	**Ingen smerte**	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	**Uudholdelige smerter**
Vurder dit niveau af energi	**Masser af energi**	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	**Ingen energi**
Vurder dit niveau af stivhed	**Ingen stivhed**	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	**Svær stivhed**
Vurder kvaliteten af din søvn	**Vågnede veludhvilet**	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	**Vågnede uudhvilet**
Vurder dit niveau af nedtrykthed	**Ikke nedtrykt**	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	**Meget nedtrykt**
Vurder dit niveau af hukommelsesproblemer	**God hukommelse**	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	**Meget dårlig hukommelse**
Vurder dit niveau af angst	**Ingen angst**	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	**Megen angst**
Vurder dit niveau af ømhed ved berøring	**Ingen ømhed**	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	**Megen ømhed**
Vurder dit niveau af balanceproblemer	**Ingen balance-problemer**	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	**Svære balance-problemer**
Vurder dit niveau af overfølsomhed for høje lyde, skarpt lys, lugte og kulde	**Ikke overfølsom**	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	**Extrem overfølsom**

Sub-total ***(kun til internt brug)*** ☐

FIQR TOTAL ***(kun til internt brug)*** ☐

Appendix B

FSQ

ID: | | | | | | | | | | |

Dags dato: dag | | / måned | | / årstal | | | | |

1. Sæt kryds ved alle de områder, hvor du har haft smerter inden for den sidste uge:

☐ Skulder, venstre ☐ Skulder, højre	☐ Øvre del af ben, venstre ☐ Øvre del af ben, højre	☐ Lænderyg ☐ Øvre del af ryggen ☐ Nakke
☐ Hofte, venstre ☐ Hofte, højre	☐ Nedre del af ben, venstre ☐ Nedre del af ben, højre	
☐ Øvre del af arm, venstre ☐ Øvre del af arm, højre	☐ Kæbe, venstre ☐ Kæbe, højre	☐ Ingen smerte i nogen af disse områder
☐ Nedre del af arme, venstre ☐ Nedre del af arm, højre	☐ Brystkasse ☐ Mave	

2. Angiv sværhedsgraden for hvert af de 3 nedenstående symptomer <u>i løbet af den sidste uge</u> efter følgende skala:

0 = Ingen problemer
1 = Lette problemer: Ubetydelige eller forbigående
2 = Moderate problemer: Betydelige, ofte forekommende og/eller på et moderat niveau
3 = Svære problemer: Udtalte, vedvarende, livs-forstyrrende

(sæt èt kryds for hvert af de tre symptomer)

	Ingen problemer	Lette problemer	Moderate problemer	Svære problemer
Træthed				
Vågner <u>u</u>-udhvilet				
Problemer med f.eks. hukommelses- og koncentrationsbesvær				

3. Har du <u>i løbet af de sidste 6 måneder</u> haft nogle af de følgende symptomer?

(sæt ét kryds for hvert af de tre symptomer)

	Ja	Nej
Hovedpine		
Smerter eller kramper i nedre del af maven		
Nedtrykhed		

Clinical Pain Research

Karin Due Bruun*, Hanne Irene Jensen, Morten Rune Blichfeldt-Eckhardt, Henrik Bjarke Vaegter, Palle Toft, Kirstine Amris and Nina Kvorning

Performance of the 2016 diagnostic criteria for fibromyalgia in a tertiary care pain rehabilitation setting: a diagnostic accuracy study

https://doi.org/10.1515/sjpain-2021-0081
Received May 7, 2021; accepted October 5, 2021;
published online October 20, 2021

Abstract

Objectives: With the International Classification of Diseases 11th revision (classifying fibromyalgia as a primary pain disorder) soon to be implemented, the importance of pain physicians being able to identify patients with fibromyalgia is emphasized. The diagnostic criteria proposed in 2016 are based on self-reported pain distribution and symptom severity. The study aimed to evaluate the diagnostic accuracy of the 2016 diagnostic criteria for fibromyalgia applied in a population of patients with high impact chronic pain referred for pain rehabilitation.
Methods: The study was performed as a diagnostic accuracy study at two Danish interdisciplinary pain rehabilitation centers, including 215 participants. All participants were evaluated clinically to identify patients with fibromyalgia. The diagnosis was based on expert opinion, but the minimum requirements were: (1) pain in all four body quadrants and axially for at least three months and (2) minimum 8 of 18 positive tender points. Participants filled in the fibromyalgia survey questionnaire, the patient version of the 2016 diagnostic criteria. Sensitivity, specificity, likelihood ratios, and positive and negative post-test probabilities were calculated using a clinical diagnosis of fibromyalgia as the reference standard.
Results: Based on clinical diagnosis 45% of the participants were diagnosed with fibromyalgia; of these, only 19% had been diagnosed previously. The 2016 diagnostic criteria demonstrated a sensitivity of 88.5%, a specificity of 81.5%, a positive likelihood ratio of 4.79, a negative likelihood ratio of 0.14, a positive post-test probability of 79.4%, and a negative post-test probability of 10.2%.
Conclusions: Fibromyalgia was severely under-diagnosed among patients with high impact chronic pain referred to tertiary care in two pain rehabilitation centers in Denmark. The 2016 diagnostic criteria showed sufficient discriminatory properties suggesting that the fibromyalgia survey questionnaire can be used as a screening tool assisting the identification of fibromyalgia in this patient population.

Keywords: diagnostic accuracy; diagnostic criteria; fibromyalgia; sensitivity; specificity.

***Corresponding author: Karin Due Bruun**, Pain Centre, Pain Research Group, University Hospital Odense, Heden 7-9, Indgang 200, DK – 5000, Odense, Denmark, Phone: +45 26183619, Fax: +45 65415064, E-mail: karin.due.bruun@rsyd.dk. https://orcid.org/0000-0002-3106-2461
Hanne Irene Jensen, Department of Anesthesiology and Intensive Care, Vejle and Middelfart Hospitals, University Hospital of Southern Denmark, Odense, Denmark; and Department of Regional Health Research, University of Southern Denmark, Odense, Denmark. https://orcid.org/0000-0001-9323-4284
Morten Rune Blichfeldt-Eckhardt and Henrik Bjarke Vaegter, Pain Centre, Pain Research Group, Odense University Hospital, Odense, Denmark; and Department of Clinical Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark. https://orcid.org/0000-0002-7707-9947 (H.B. Vaegter)
Palle Toft, Department of Anesthesiology and Intensive Care, Odense University Hospital, Odense, Denmark
Kirstine Amris, Department of Rheumatology, The Parker Institute, Copenhagen University Hospital, Frederiksberg, Denmark
Nina Kvorning, Pain Centre, Pain Research Group, Odense University Hospital, Odense, Denmark; and Department of Anaesthesiology, Multidisciplinary Pain Centre, Vejle and Middelfart Hospitals, University Hospital of Southern, Odense, Denmark

Introduction

Fibromyalgia (FM) is a pain disorder characterized by chronic widespread pain (CWP) and generalized mechanical hyperalgesia. Other significant symptoms include insomnia, fatigue, and cognitive dysfunction [1]. Patients with FM represent a subgroup of CWP presenting with a higher symptom burden and lower functional ability [2, 3]. In Denmark, a rheumatologist traditionally diagnoses patients with FM, but daily care is provided by primary care practitioners, who can choose to refer severely affected

patients (high-impact chronic pain [4–7]) to a tertiary care pain rehabilitation center. When the new International Classification of Diseases 11th revision (ICD-11) is implemented in 2022, fibromyalgia will be classified as a primary pain disorder (a subtype of CWP) instead of a rheumatological disease [8]. Consequently, in the future, specialists working with chronic non-malignant pain (CNMP) patients should be able to identify individuals with fibromyalgia.

No objective biological markers can validate a diagnosis of fibromyalgia, and several criteria for the classification and diagnosis of fibromyalgia have been suggested over the years. The purpose of classification criteria is to identify well-defined homogeneous groups of patients for research purposes, requiring high specificity. In contrast, diagnostic criteria strive to capture as many cases as possible and must perform with high sensitivity [9]. However, the utilization of diagnostic tests in patient care settings must be guided by evidence, and predictive values are greatly influenced by the prevalence of the disease and should not be generalized beyond the studied population.

The most widely used fibromyalgia research criteria are the dual 1990 American College of Rheumatology (ACR) classification criteria requiring both anamnestic widespread pain and widespread mechanical hyperalgesia assessed by manual tender point examination [10]. When developing these criteria, tender point count was the most powerful discriminator separating fibromyalgia from other painful rheumatic disorders [10]. The ACR 1990 uses 11 positive tender points as cut-off, but later studies have shown that a shift in disease severity in patients with CWP occurs at eight tender points [11]. In the clinical context, the ACR1990 classification criteria have been criticized for placing the diagnosis at the far end of a severity spectrum and ignoring other important key symptoms [12, 13]. This led to the proposal of diagnostic symptom-based criteria in 2010, and the definition of fibromyalgia expanded to include symptoms other than pain [14]. For the 2010 diagnostic criteria, three pain sites were sufficient for the definition. With a revision of the criteria in 2016 a generalized pain criterion was added, requiring pain in four out of five body regions with a minimum of four pain sites [15]; these criteria are referred to as the 2016 diagnostic criteria. The 2016 diagnostic criteria have been validated in populations with rheumatic disorders [16–18] and in a Norwegian population of CNMP patients [19], showing acceptable internal consistency and good construct validity.

To our knowledge, only one study has investigated the diagnostic accuracy of the 2016 diagnostic criteria in a tertiary care pain rehabilitation setting, using the ACR1990 criteria as the reference standard [20]. Only patients with CWP were included, and the 2016 diagnostic criteria demonstrated both low sensitivity and specificity. However, as diagnostic and classification criteria have been developed for different purposes, using classification criteria as a reference standard for the diagnostic accuracy of diagnostic criteria is questionable.

This study aimed to investigate the diagnostic accuracy of the 2016 diagnostic criteria for fibromyalgia in a cross-sectional sample of patients with mixed CNMP disorders referred to tertiary care, using a clinical diagnosis of fibromyalgia by expert opinion as the reference standard.

Materials and methods

Study design

The study was designed as a prospective diagnostic accuracy study. The performance of the 2016 diagnostic criteria for fibromyalgia was investigated in a cross-sectional sample of patients with mixed chronic pain conditions referred to tertiary care. The reference standard was a clinical diagnosis of fibromyalgia by expert opinion. The study was performed according to Standards for Reporting Diagnostic accuracy studies (STARD) 2015 guidelines [21].

Participants

Participants were recruited from two Danish public interdisciplinary pain rehabilitation centers. One was a university hospital unit (Center 1) and the other a general hospital unit (Center 2). Both centers received patients with different CNMP disorders referred by primary care practitioners. Patients referred to pain rehabilitation in Denmark must be sufficiently examined for diseases accessible to causal treatment before referral. Most patients referred will suffer from primary chronic pain conditions displaying nociplastic pain features [22]. Only a minor group consists of patients with rheumatological or neurological diseases, where disabling pain is still present despite successful causal treatment.

A specific chief consultant at each pain center recruited the study participants. From December 10th, 2017, patients were invited to participate at their first appointment with one of the two responsible consultants for two consecutive years. As one of the consultants changed employment during the study period, inclusion was terminated earlier in Center 2 (end of August 2019). Inclusion criteria were patients aged 18 years or above who could read and understand Danish. Patients who were not able to complete questionnaires due to poor somatic or mental status were excluded.

The 2016 diagnostic criteria

The 2016 diagnostic criteria consist of a Widespread Pain Index (WPI), a Symptom Severity Score (SSS), and a generalized pain criterion [15]. The questionnaire can be completed by a physician or patient. The patient version used in this study is referred to as the fibromyalgia survey questionnaire (FSQ) [23]. To fulfill the 2016 criteria, the patient had to have a WPI of at least seven and a SSS of at least five *or* a WPI of 4–6 and a SSS of at least nine. Furthermore, a generalized pain

criterion had to be satisfied, defined as pain in a minimum of four of five regions.

Translation procedure

The FSQ was translated to Danish according to standardized guidelines [24]. The questionnaire was initially translated independently from English to Danish by a panel of two researchers and one nonprofessional. Based on these three translations, the panel agreed on a Danish version. A new panel of two researchers and one nonprofessional then translated the Danish version back to English. Finally, all six members of the two panels collectively evaluated the Danish version and the reverse English translation and agreed on a final Danish version. Members of both panels had to be fluent in both Danish and English. The final version was pilot tested in six patients with FM or other CNMP conditions. The patients were asked to fill in the questionnaire and were consecutively interviewed regarding comprehensibility. Approval was obtained from the developers to use and translate the instrument. The Danish version is presented in the Supplementary Material.

Assessment of widespread pain and tender point examination

Information about pain distribution was obtained both from interviews and pain drawings (body chart). In this study, CWP was defined as persistent or recurrent pain during the past three months located both axially and in all four body quadrants (representing five pain regions). The presence of widespread tenderness was evaluated based on examination of the 18 tender points defined by the ACR1990 criteria [10]. The tender point examination was carried out using a handheld spring-based pressure algometer with a probe area of 1 cm^2. At each point, increasing pressure was slowly applied until a maximum of 4 kg/cm^2. If a pain reaction was observed (vocalized, grimace or flinch) before or at a pressure of 4 kg/cm^2, the point was considered a positive tender point. The two consultants trained the procedure together before and during the trial to secure a high and identical quality of testing [25]. Consultants were blinded to the responses on the fibromyalgia survey during the interview and the tender point examination.

Reference standard

The reference standard was a clinical diagnosis of FM by expert opinion based on full history and clinical examination. To be diagnosed with FM the patients had to fulfill the criteria for CWP as described above. Furthermore, they were required to have symptoms of widespread pressure hyperalgesia, demanding a minimum of eight positive tender points (moderate pain = vocalized, grimace, or flinch) located both over and below the waist.

Both consultants had long-term experience in evaluating pain patients and received training from a senior rheumatologist and fibromyalgia expert to ensure a uniform evaluation of the clinical diagnose of fibromyalgia. Continuous consensus in diagnosing was made between the two pain specialists throughout the study, and in complicated cases, the senior rheumatologist was consulted.

Other assessments

Demographic data were collected about age, gender; body mass index, work status, medication, comorbidity, and existing FM diagnosis. Comorbid diseases were required to be in a stable and inactive state. Disease activity was evaluated using information from the patients' file only, and we did not perform new laboratory testing. To evaluate the level of pain and the intensity of common pain-related symptoms, the participants were asked to complete the symptom domain of the symptom impact questionnaire-revised (SIQR) [26, 27].

Statistical analyses

Our study group previously estimated the prevalence of FM in the target population to be 37% based on the 2016 diagnostic criteria [28]. However, the diagnostic accuracy of the 2016 criteria was unknown, and we suspected the true prevalence to be lower. With an estimated prevalence of FM set at 30%, a minimum sample size of 103 subjects (including 31 subjects with FM) would be required to achieve a minimum power of 80% to detect a change in the percentage value of sensitivity from 70.0 to 90.0 based on a target significance level of 0.05 [29]. A minimum sample of 44 subjects (including 13 subjects with FM) is required to detect a change in the percentage value of specificity from 70.0 to 90.0. As the prevalence estimate was uncertain, it was decided to include patients over two consecutive years, approximately 200 patients.

Data were collected in paper format, entered into a database, and transferred to the statistical program Stata 16 for analysis. Descriptive statistics were used to describe the population, using numbers and percent to present categorical variables and median and interquartile ranges to present continuous and numerical variables. Comparative analyses were performed to investigate any differences in demographic data, pain characteristics, and level of pain-related symptoms between patients with and without FM based on the clinical diagnosis. Comparative analyses were also made between FM patients with and without an existing diagnose. The chi-square test, Fisher's exact test, or the Wilcoxon rank-sum test was used according to data type. A p-value < 0.05 was considered significant.

To analyze performance characteristics of the 2016 diagnostic criteria, contingency tables were made using the clinical diagnosis as the reference standard. Data were derived from the tables to calculate sensitivity, specificity, and positive and negative likelihood ratios. Furthermore, we have examined the positive post-test probability (PPTP) and the negative post-test probability (NPTP). PPTP is the probability of having the disease if the test is positive. PPTP has the same value as the positive predictive value (PPV). The negative post-test probability (NPTP) is the probability of having the disease if the test is negative. NPTP equals one minus the negative predictive value (NPV).

Results

Study participation

Of the 297 patients screened for eligibility, 275 met the inclusion criteria, and 215 patients were included in the

study, giving an overall participation rate of 78% (Figure 1). Of these, 163 (76%) were included at the university hospital unit and 52 (24%) at the general hospital unit, with 76 and 87% participation rates, respectively. Patient characteristics are presented in Table 1.

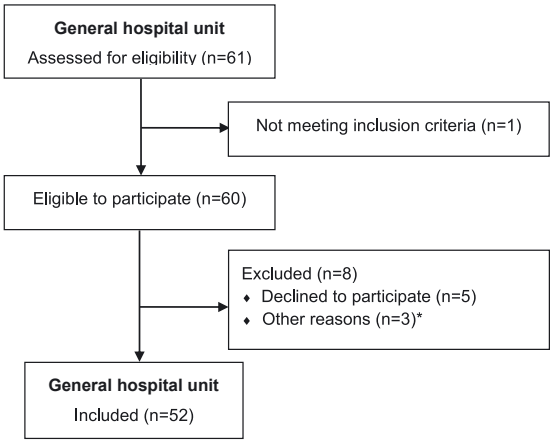
Diagnose of FM based on expert opinion

Among the 215 participants evaluated by the two expert pain physicians, 96 were diagnosed with FM equivalent with 45% of this sample (Table 2). In this study population, 15% of the men and 55% of the women were diagnosed with FM.

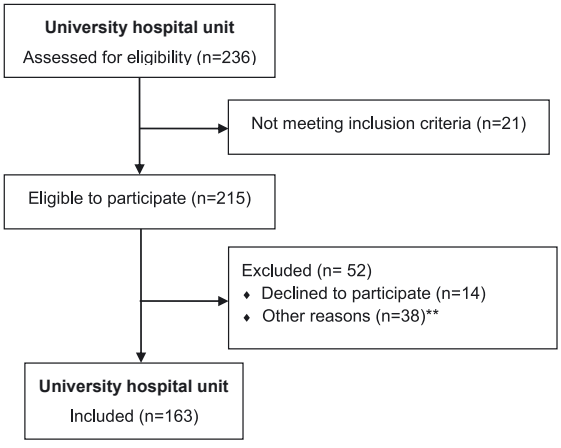
Performance of the 2016 diagnostic criteria

The contingency table and performance characteristics for the 2016 diagnostic criteria using a clinical diagnosis as the reference standard are presented in Tables 2 and 3. The criteria demonstrated a sensitivity of 88.5%, a specificity of 81.5%, a positive likelihood ratio of 4.79, a negative likelihood ratio of 0.14, a positive post-test probability of 79.4%, and a negative post-test probability of 10.2%.

Contingency table and performance characteristics showing how the criteria would perform if the 1990ACR classification criteria had been used as the reference standard instead are presented in Supplementary Tables 1 and 2.



*Other reasons: treatment ended before second visit (n=2), excluded because of very poor somatic or mental status (n=1)



**Other reasons: treatment ended before second visit (n=27), excluded because of very poor somatic or mental status (n=11)

Figure 1: Overview of participant flow. *Other reasons: Treatment ended before the second visit (n=2); excluded due to very poor somatic or mental status (n=1). **Other reasons: Treatment ended before the second visit (n=27); excluded due to very poor somatic or mental status (n=11).

Table 1: Patient characteristics showing the distribution for the entire cohort and each pain center.

Variable	Total (n=215)	Centre 1 (n=163)	Centre 2 (n=52)
Women n, %	160 (74)	123 (76)	37 (71)
Age in years (median, quartiles)	49 (38, 58)	50 (38, 58)	45 (38, 53.5)
BMI kg/m^2 (median, quartiles)	26.8 (24, 32)	28.2 (24, 33)	25.6 (23, 29)
Pain duration in years (median, quartiles)	10 (5, 20)	11 (6, 20)	7 (4, 19)
Employment situation [a]:			
Working on ordinary conditions n, %	30 (14)	18 (11)	12 (23)
Subsidized job n, %	46 (21)	40 (25)	6 (12)
Sick leave n, %	52 (24)	40 (25)	12 (23)
Disability pension n, %	40 (19)	30 (18)	10 (19)
Retired n, %	23 (11)	17 (10)	6 (12)
Studying n, %	15 (7)	13 (8)	2 (4)
Fibromyalgia diagnosed before referral n, %	40 (19)	32 (20)	8 (16)
Fibromyalgia diagnosed after referral n, %	96 (45)	73 (45)	23 (44)

[a]Different n due to missing data.

Table 2: 2 × 2 contingency table for the 2016 diagnostic criteria vs. a clinical diagnose of fibromyalgia based on expert opinion.

2016 diagnostic criteria	Clinical diagnose of fibromyalgia		Total
	Yes	No	
Positive	85	22	107
Negative	11	97	108
Total	96	119	215

Differences in characteristics between patients with and without FM

Differences regarding demographics and pain characteristics are shown in Tables 4 and 5. There were significantly more women in the FM group, and FM patients were more often in a subsidized job, on sick leave, or disability pension. Longer duration of pain, higher intensity of both pain, and all other SIQR items, including the symptom domain score, were observed for the FM group. FM patients used opioids less frequently and low dose naltrexone more regularly.

Differences in characteristics between FM patients with and without an existing diagnose

Among the 96 participants diagnosed with FM in this study, 40 (19%) had been diagnosed before referral (Table 1). There were no men with an existing FM diagnosis, but eight male FM patients were identified in the study. The group with an existing FM diagnose (FM-E) was compared to the group diagnosed only in the study (FM-S) (Tables 4 and 5). The number of tender points was significantly higher in the FM-E group. Otherwise, no differences were found regarding demographics, pain characteristics, or level of symptoms. The number of patients who fulfilled the ACR1990 criteria was comparable between the FM-E and FM-S groups.

The expert pain physicians confirmed all patients with an existing FM diagnosis to have FM. Among the 56 patients from the FM-S group, the referral diagnoses were CWP (n=15), rheumatological disease (n=6), back pain (n=13), neck pain (n=7), other localized pain (n=7), or unspecified pain (n=8).

Table 3: Performance characteristics of the 2016 diagnostic criteria for fibromyalgia, using a clinical diagnosis as the reference standard.

Sensitivity	Specificity	Positive likelihood ratio	Negative likelihood ratio	Positive post-test probability	Negative post-test probability
0.885 (0.822–0.949)[a]	0.815 (0.745–0.885)[a]	4.789 (3.262–7.032)[a]	0.141 (0.080–0.247)[a]	0.794 (0.718–0.871)[a]	0.102 (0.045–0.159)[a]

[a]95% confidence interval.

Table 4: Differences in demographic data between patients with and without fibromyalgia and between patients with fibromyalgia diagnosed before and after referral.

Variable	No fibromyalgia (n=119)	Fibromyalgia (n=96)	p-Value[a]	Fibromyalgia previously diagnosed (n=40)	Fibromyalgia not previously diagnosed (n=56)	p-Value[a]
Women n, %	72 (61)	88 (92)	<0.001	40 (100)	48 (85)	<0.05
Age in years (median, quartiles)	49 (37, 61)	48.5 (41, 55)	0.42	49.5 (43, 55.5)	48.5 (37.5, 53)	0.23
Pain duration in years (median, quartiles)	8 (4, 17)	13 (8, 21)	<0.001	15 (9, 25)	11 (6, 20)	0.06
Employment status:[b]			<0.05			0.12
Working on ordinary conditions n, %	22 (18)	8 (8)		4 (11)	4 (7)	
Subsidized job n, %	22 (18)	24 (25)		11 (30)	13 (23)	
Sick leave n, %	26 (22)	26 (27)		6 (16)	20 (36)	
Disability pension n, %	19 (16)	21 (22)		13 (35)	8 (14)	
Retired n, %	19 (16)	4 (4)		1 (3)	3 (5)	
Studying n, %	8 (7)	7 (7)		2 (5)	5 (9)	
Comorbidity:						
Inflammatory rheumatic disease n, %	7 (6)	9 (9)	0.33	2 (5)	7 (13)	0.21
Polyneuropathy or multiple sclerosis n, %	9 (8)	1 (1)	<0.05	0 (0)	1 (2)	0.40
Pain medication:						
Paracetamol n, %	41 (34)	26 (27)	0.25	12 (30)	14 (25)	0.59
NSAIDs n, %	13 (11)	12 (13)	0.72	5 (13)	7 (13)	1.00
Tricyclic antidepressants n, %	18 (15)	11 (11)	0.43	5 (13)	6 (11)	0.79
Gabapentin n, %	19 (16)	10 (10)	0.24	4 (10)	6 (11)	0.91
Pregabalin n, %	15 (13)	8 (8)	0.31	4 (10)	4 (7)	0.62
SNRIs n, %	24 (20)	20 (21)	0.90	11 (28)	9 (16)	0.17
Low dose naltrexone n, %	5 (4)	13 (14)	<0.05	8 (20)	5 (9)	0.12
Opioids n, %	43 (36)	22 (23)	<0.05	8 (20)	14 (25)	0.57

[a]Chi-square test or Fisher's Exact test for categorical data and Wilcoxon rank-sum test for continuous or numerical data. [b]Different n due to missing data.

Discussion

To our knowledge, this is the first study to investigate the diagnostic accuracy of the 2016 diagnostic criteria for fibromyalgia using clinical diagnosis as the reference standard in a CNMP tertiary care pain rehabilitation setting. The 2016 diagnostic criteria were translated to Danish according to standardized guidelines [24], and the Danish version is provided in the Supplementary Material and is freely available for other purposes. The 2016 diagnostic criteria showed acceptable discriminatory properties for the identification of fibromyalgia in the population examined.

Diagnosing FM in the clinic

There is an ongoing debate whether to regard fibromyalgia as a primary pain disorder or as part of a functional somatic syndrome [8, 30, 31]. However, as experts working with pain rehabilitation, pain medicine, and pain research (e.g. clinical trials investigating the effect of new pain treatments), we welcome the new ICD-11 revision, which acknowledges widespread pain to be the core symptom in fibromyalgia. The presence of widespread pain is a key symptom in both the ACR1990 classification criteria, the 2016 diagnostic criteria, and other later proposed diagnostic criteria [32–35]. Some criteria have focused on mechanical hyperalgesia as another important key symptom. In contrast, others focus more on the level of somatic symptoms. A diagnosis of FM in the clinical setting will always be based on an expert's opinion, taking symptoms, objective signs, comorbidity, medication, and the ruling out of other possible causes of the symptoms into consideration. However, a systematic approach using well-validated diagnostic criteria can be a valuable aid in identifying patients with FM.

Table 5: Differences in pain characteristics and level of symptoms between patients with and without fibromyalgia and between patients with fibromyalgia diagnosed before and after referral.

Variable	No fibromyalgia (n=119)	Fibromyalgia (n=96)	p-Value[a]	Fibromyalgia previously diagnosed (n=40)	Fibromyalgia not previously diagnosed (n=56)	p-Value[a]
Widespread pain index (WPI)[b] (median, quartiles)	5 (3, 7)	11 (9, 15)	<0.001	10.5 (9, 16)	12 (9, 14)	0.89
Symptom severity score (SSS)[c] (median, quartiles)	7 (4, 9)	10 (8, 11)	<0.001	10 (8, 11)	10 (9, 11)	0.32
Tender point count[d] (median, quartiles)	6 (2, 11)	16 (13, 18)	<0.001	17 (15, 18)	15.5 (12, 18)	<0.05
Full fills ACR1990 criteria[e] n, %	0 (0)	88 (92)	<0.001	37 (93)	51 (91)	0.80
Items from the SIQR symptom domain[f]						
Pain (median, quartiles)	6 (5, 8)	7 (6, 8)	<0.05	7 (6, 8)	7 (6, 8)	0.69
Energy (median, quartiles)	6 (4, 8)	7 (5, 8)	<0.05	7 (5, 8)	7 (5, 8.5)	0.67
Stiffness (median, quartiles)	5 (2, 7)	7 (5, 8)	<0.001	7 (5, 8)	7 (5, 8)	0.54
Sleep quality (median quartiles)	6 (2, 8)	8 (6, 9)	<0.001	8.5 (6.5, 10)	8 (6, 9)	0.38
Depression (median quartiles)	2 (0, 5)	5 (1, 7)	<0.001	5 (1, 7)	5 (1.5, 6.5)	0.99
Memory problems (median, quartiles)	5 (2, 7)	7 (5, 8.5)	<0.001	7 (6, 8)	7 (5, 9)	0.72
Anxiety (median quartiles)	0 (0, 2)	2 (0, 5)	<0.001	1 (0, 3.5)	3 (0, 6.5)	0.09
Tenderness to touch (median, quartiles)	4 (0, 7)	7 (5.5, 9)	<0.001	7 (5, 9)	7 (6, 9)	0.62
Balance problems (median, quartiles)	2 (0, 5)	5 (2, 7)	<0.001	5 (3, 7)	5 (1, 7)	0.39
Sensitivity to sensory inputs (median, quartiles)	2 (0, 7)	7 (4.5, 8)	<0.001	8 (5, 9)	7 (3, 8)	0.12
FIQR symptom domain sum score[g] (median, quartiles)	38 (28, 51)	61 (49, 70.5)	<0.001	60.5 (50.5, 69)	61 (48.5, 71)	0.94

[a]Chi-square test or Fisher's Exact test for categorical data and Wilcoxon rank-sum test for continuous or numerical data. [b]Widespread Pain Index (WPI) from the 2016 diagnostic criteria. The WPI counts the number of painful body parts during the past 7 days, giving rise to a score between 0 and 19. [c]Symptom Severity Score (SSS) from the 2016 diagnostic criteria. SSS is ranging from 0–12. [d]Tender point examination was carried out using a handheld pressure algometer. Each tender point was examined applying a slowly increased pressure until 4 kg/cm^2. If the threshold for pain was reached, the tender point was considered positive. The number of positive tender points is ranging from 0–18. [e]All participants were classified according to the American College of Rheumatology classification criteria from 1990 (ACR1990). The participants full filled the criteria if they reported having had widespread pain (pain in all four body quadrant plus axially) during the last 3 months and had a minimum of 11 out of 18 positive tender points. [f]Level of pain and the intensity of nine common pain-related symptoms, were evaluated using the 10 items from the symptom part of the Symptom Impact Questionnaire-Revised (SIQR). Each item evaluates the average severity of the symptom during the last 7 days on a 0–10 numeric rating scale (NRS), 0 indicating "no problem" and 10 indicating "severe problems." [g]The SIQR symptom domain sum score is calculated as the sum of the 10 items (range 0–100).

Diagnosing FM in the study population

In this study, we diagnosed FM based on expert opinion, considering both full history and clinical examination. To secure homogeneity and transparency of the diagnosis, some minimum requirements were agreed on (1) pain located in all four body quadrants plus axially for at least three months, and (2) the presence of a minimum 8 of 18 positive tender points. This was based on a previous study showing that a shift in severity of disease in CWP patients occurs at eight positive tender points [11]. The tender point examination is a quick, easy to learn, and reliable instrument in the clinic [25, 36]. Most of the patients diagnosed with FM in the study also fulfilled the ACR1990 criteria.

That fibromyalgia is more prevalent among women, with a male to female ratio of 1:3 found in the general population [37], was observed in our present sample with a ratio of 1:3.7. Only 19% of the patients diagnosed with FM had an established diagnosis, and notably, none of them were men. Patients diagnosed with FM only in the study did not differ from patients with an existing FM diagnosis, except they had a lower tender point count.

In our study population, 45% were diagnosed with FM based on expert opinion, 41% fulfilled the ACR1990 criteria, and 50% fulfilled the ACR2016 criteria. In a previous study, including 1,343 patients from Center 1, only 37% fulfilled the ACR2016 criteria [28]. The demographic characteristics for the two study populations are comparable except regarding gender, showing a higher ratio of women in our present sample (74 vs. 66%), which could explain the observed differences in FM frequency. In a recent study including more than 12.000 patients from all tertiary pain centers in Denmark, 68% were women [22].

Performance of the 2016 diagnostic criteria

Even though the diagnostic accuracy of the 2016 diagnostic criteria has previously been shown to be acceptable in the general population and populations of patients with different rheumatic diseases [15], they might perform differently in CNMP populations. In our mixed CNMP tertiary care sample, the 2016 diagnostic criteria demonstrated high sensitivity and acceptable specificity. Thus, in this study population, the 2016 diagnostic criteria showed sufficient discriminatory properties, indicating that the FSQ can be used as an easily applied tool to assist systematic screening for FM in patients referred for tertiary pain rehabilitation.

In a previous diagnostic accuracy study of the 2016 diagnostic criteria in a tertiary care pain rehabilitation setting, only patients with CWP were included. Using the ACR1990 criteria as reference standard they found the 2016 criteria to perform poorly [20]. This finding suggests that the 2016 diagnostic criteria might not be helpful in discriminating FM from CWP with lower severity in populations with high impact pain.

Implications of diagnosing FM

It is well known that patients with fibromyalgia have a higher symptom burden and lower functional ability compared to patients with other chronic pain disorders, including patients with CWP who do not fulfill the ACR1990 criteria [3, 38]. In accordance with this, we found that patients diagnosed with FM had a significantly longer duration of pain, higher symptom burden, and were more work incapacitated than the non-FM group.

We found FM to be severely under-diagnosed in this cross-sectional sample of patients with mixed high-impact chronic pain conditions referred to pain rehabilitation. Only 15 of the 56 patients not previously recognized to have FM were referred with a diagnosis of CWP, whereas the remaining 41 were referred with localized or unspecified pain diagnoses. Despite rapidly increasing insight into pathophysiological mechanisms [39, 40], fibromyalgia remains a controversial diagnosis due to limited knowledge about the disease among many healthcare professionals [41, 42], and the diagnosis is often delayed by several years [43]. A correct and timely diagnosis is of great value to patients as it encourages better self-management [44] and helps healthcare providers to offer appropriate care [42]. We suggest that a systematic approach, including the use of the FSQ, could easily be implemented as a routine to assist the identification of FM among patients with chronic pain referred to tertiary care.

Strengths and limitations

To secure an accurate reference standard and to limit interrater variability, patients were recruited by two senior pain specialists who were trained in diagnosing fibromyalgia and performing tender point examination. Secretaries otherwise not involved in the study allocated all referred patients randomly to all the specialists working at the centers, striving to secure a valid cross-sectional sample of the total population from the two pain centers. However, this cross-sectional sample had a high frequency of women, explaining the relatively high frequency of FM found in the present study. With a fibromyalgia prevalence of 45% in this population, the sample size was sufficient to estimate the diagnostic accuracy with acceptable power [29]. To reduce bias the investigators were blinded regarding the FSQ and the participants were not informed whether they were diagnosed with FM in the study until after they had completed the FSQ.

As the study was performed at two pain centers in Denmark, our results might not be generalizable to other tertiary care chronic pain populations in Denmark or other countries. In Denmark, all patients referred to pain rehabilitation receive the diagnose “chronic complex non-malignant pain.” We did not collect data about the history of trauma, surgery, or previous cancer treatment in the present study. The implementation of the ICD-11 will secure diagnosis of primary and secondary pain in future trials.

Conclusions

The 2016 diagnostic criteria showed high sensitivity and acceptable specificity in the study population and were able to capture the spectrum of the disease, but also included some false positives. Our results suggest that the FSQ might be useful as an easily applied screening tool to

assist the identification of patients with fibromyalgia in tertiary care chronic pain settings. However, more studies investigating the diagnostic accuracy of the 2016 criteria in larger CNMP populations and other settings and countries are warranted.

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.
Research funding: The authors state no funding is involved.
Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: The authors state no conflicts of interest.
Informed consent: Informed consent was obtained from all individuals included in this study.
Ethical approval: The study was performed according to the Declaration of Helsinki. Approval was obtained from the Danish Data Protection Agency (ref. no. 17/35274). According to Danish law, approval from The Regional Committees on Health Research Ethics was not needed, and this was confirmed by the Ethics Committees (01.02.2017).

References

1. Jay GW, Barkin RL. Fibromyalgia. Dis Mon 2015;61:66–111.
2. Coster L, Kendall S, Gerdle B, Henriksson C, Henriksson KG, Bengtsson A. Chronic widespread musculoskeletal pain – a comparison of those who meet criteria for fibromyalgia and those who do not. Eur J Pain 2008;12:600–10.
3. Amris K, Waehrens EE, Jespersen A, Bliddal H, Danneskiold-Samsoe B. Observation-based assessment of functional ability in patients with chronic widespread pain: a cross-sectional study. Pain 2011;152:2470–6.
4. Dahlhamer J, Lucas J, Zelaya C, Nahin R, Mackey S, DeBar L, et al. Prevalence of chronic pain and high-impact chronic pain among adults – United States, 2016. MMWR Morb Mortal Wkly Rep 2018; 67:1001–6.
5. Zelaya CE, Dahlhamer JM, Lucas JW, Connor EM. Chronic pain and high-impact chronic pain among U.S. Adults. NCHS Data Brief 2019;2020:1–8.
6. Von Korff M, Scher AI, Helmick C, Carter-Pokras O, Dodick DW, Goulet J, et al. United States national pain strategy for population research: concepts, definitions, and pilot data. J Pain 2016;17: 1068–80.
7. Vaegter HB, Christoffersen LO, Enggaard TP, Holdggard DEM, Lefevre TN, Eltved R, et al. Socio-demographics, pain characteristics, quality of life and treatment values before and after specialized interdisciplinary pain treatment: results from the Danish clinical pain registry (PainData). J Pain Res 2021;14:1215–30.
8. Nicholas M, Vlaeyen JWS, Rief W, Barke A, Aziz Q, Benoliel R, et al. The IASP classification of chronic pain for ICD-11: chronic primary pain. Pain 2019;160:28–37.
9. Aggarwal R, Ringold S, Khanna D, Neogi T, Johnson SR, Miller A, et al. Distinctions between diagnostic and classification criteria? Arthritis Care Res 2015;67:891–7.
10. Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, 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.
11. Amris K, Waehrens EE, Jespersen A, Stockmarr A, Bennett R, Bliddal H, et al. The relationship between mechanical hyperalgesia assessed by manual tender point examination and disease severity in patients with chronic widespread pain: a cross-sectional study. Int J Rheumatol 2014;2014:417596.
12. Wolfe F. The relation between tender points and fibromyalgia symptom variables: evidence that fibromyalgia is not a discrete disorder in the clinic. Ann Rheum Dis 1997;56:268–71.
13. Wolfe F. Stop using the American college of rheumatology criteria in the clinic. J Rheumatol 2003;30:1671–2.
14. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Katz RS, Mease P, et al. The American college of rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res 2010;62:600–10.
15. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Hauser W, Katz RL, et al. Revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum 2016;46:319–29.
16. Shresher NM, Mohamed AE, Elshahaly MH. Performance of 2016 revised fibromyalgia diagnostic criteria in patients with rheumatoid arthritis. Rheumatol Int 2019;39:1703–10.
17. Kang JH, An M, Choi SE, Xu H, Park DJ, Lee JK, et al. Performance of the revised 2016 fibromyalgia diagnostic criteria in Korean patients with fibromyalgia. Int J Rheum Dis 2019;22:1734–40.
18. Ablin JN, Wolfe F. A comparative evaluation of the 2011 and 2016 criteria for fibromyalgia. J Rheumatol 2017;44:1271–6.
19. Fors EA, Wensaas KA, Eide H, Jaatun EA, Clauw DJ, Wolfe F, et al. Fibromyalgia 2016 criteria and assessments: comprehensive validation in a Norwegian population. Scand J Pain 2020;20:663–72.
20. Ahmed S, Aggarwal A, Lawrence A. Performance of the American college of rheumatology 2016 criteria for fibromyalgia in a referral care setting. Rheumatol Int 2019;39:1397–403.
21. Korevaar DA, Cohen JF, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, et al. Updating standards for reporting diagnostic accuracy: the development of STARD 2015. Res Integr Peer Rev 2016;1:7.
22. Vaegter HB, Hoybye MT, Bergen FH, Parsons CE. Sleep disturbance in patients attending specialized chronic pain clinics in Denmark: a longitudinal study examining the relationship between sleep and pain outcomes. Scand J Pain 2021;21: 539–47.
23. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Hauser W, Katz RS, et al. Fibromyalgia criteria and severity scales for clinical and epidemiological studies: a modification of the ACR preliminary diagnostic criteria for fibromyalgia. J Rheumatol 2011;38: 1113–22.
24. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. Value Health 2005;8:94–104.
25. Cott A, Parkinson W, Bell MJ, Adachi J, Bedard M, Cividino A, et al. Interrater reliability of the tender point criterion for fibromyalgia. J Rheumatol 1992;19:1955–9.

26. 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.
27. Friend R, Bennett RM. Distinguishing fibromyalgia from rheumatoid arthritis and systemic lupus in clinical questionnaires: an analysis of the revised fibromyalgia impact questionnaire (FIQR) and its variant, the symptom impact questionnaire (SIQR), along with pain locations. Arthritis Res Ther 2011;13:R58.
28. Plesner KB, Vaegter HB. Symptoms of fibromyalgia according to the 2016 revised fibromyalgia criteria in chronic pain patients referred to multidisciplinary pain rehabilitation: influence on clinical and experimental pain sensitivity. J Pain 2018;19:777–86.
29. Bujang MA, Adnan TH. Requirements for minimum sample size for sensitivity and specificity analysis. J Clin Diagn Res 2016;10: YE01–YE6.
30. Henningsen P, Layer P, Fink P, Hauser W. Chronic primary pain: a pain-centered view of the world is too narrow. Pain 2019;160:1683.
31. Wolfe F. The history of the idea of widespread pain and its relation to fibromyalgia. Scand J Pain 2020;20:647–50.
32. Bennett RM, Friend R, Marcus D, Bernstein C, Han BK, Yachoui R, et al. Criteria for the diagnosis of fibromyalgia: validation of the modified 2010 preliminary American college of rheumatology criteria and the development of alternative criteria. Arthritis Care Res 2014;66:1364–73.
33. Arnold LM, Bennett RM, Crofford LJ, Dean LE, Clauw DJ, Goldenberg DL, et al. AAPT diagnostic criteria for fibromyalgia. J Pain 2019;20:611–28.
34. Salaffi F, Sarzi-Puttini P, Girolimetti R, Gasparini S, Atzeni F, Grassi W. Development and validation of the self-administered fibromyalgia assessment status: a disease-specific composite measure for evaluating treatment effect. Arthritis Res Ther 2009; 11:R125.
35. Stewart JA, Mailler-Burch S, Muller D, Studer M, von Kanel R, Grosse Holtforth M, et al. Rethinking the criteria for fibromyalgia in 2019: the ABC indicators. J Pain Res 2019;12:2115–24.
36. Jensen OK, Callesen J, Nielsen MG, Ellingsen T. Reproducibility of tender point examination in chronic low back pain patients as measured by intrarater and inter-rater reliability and agreement: a validation study. BMJ Open 2013;3. https://doi.org/10.1136/bmjopen-2012-002532.
37. Queiroz LP. Worldwide epidemiology of fibromyalgia. Curr Pain Headache Rep 2013;17:356.
38. Schaefer C, Mann R, Masters ET, Cappelleri JC, Daniel SR, Zlateva G, et al. The Comparative burden of chronic widespread pain and fibromyalgia in the United States. Pain Pract 2016;16:565–79.
39. Benlidayi IC. Role of inflammation in the pathogenesis and treatment of fibromyalgia. Rheumatol Int 2019;39:781–91.
40. Ceko M, Bushnell MC, Gracely RH. Neurobiology underlying fibromyalgia symptoms. Pain Res Treat 2012;2012:585419.
41. Hayes SM, Myhal GC, Thornton JF, Camerlain M, Jamison C, Cytryn KN, et al. Fibromyalgia and the therapeutic relationship: where uncertainty meets attitude. Pain Res Manag 2010;15:385–91.
42. Perrot S, Choy E, Petersel D, Ginovker A, Kramer E. Survey of physician experiences and perceptions about the diagnosis and treatment of fibromyalgia. BMC Health Serv Res 2012;12:356.
43. Choy E, Perrot S, Leon T, Kaplan J, Petersel D, Ginovker A, et al. A patient survey of the impact of fibromyalgia and the journey to diagnosis. BMC Health Serv Res 2010;10:102.
44. Doebl S, Macfarlane GJ, Hollick RJ. 'No one wants to look after the fibro patient'. Understanding models, and patient perspectives, of care for fibromyalgia: reviews of current evidence. Pain 2020;161:1716–25.

Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/sjpain-2021-0081).

Pain Medicine, 0(0), 2020, 1–9
doi: 10.1093/pm/pnaa001
Original Research

Low-Dose Naltrexone for the Treatment of Fibromyalgia: Investigation of Dose–Response Relationships

Karin Bruun-Plesner, MD,*,†,‡ Morten Rune Blichfeldt-Eckhardt, MD, PhD,*,† Henrik Bjarke Vaegter, MSc, PhD,*,† Joergen T. Lauridsen, MSc, PhD,§ Kirstine Amris, MD, MedScD,¶ and Palle Toft, MD, PhD, MedScD‖

*Pain Research Group, Pain Centre, Odense University Hospital, Odense, Denmark; †Department of Clinical Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark; ‡Open Patient Data Explorative Network (OPEN), Odense University Hospital, Odense, Denmark; §Department of Business and Economics, University of Southern Denmark, Odense, Denmark; ¶Department of Rheumatology, The Parker Institute, Copenhagen University Hospital, Bispebjerg-Frederiksberg, Denmark; ‖Department of Anesthesiology and Intensive Care, Odense University Hospital, Odense, Denmark

Correspondence to: Karin Bruun Plesner, Pain Research Group, Pain Center, University Hospital Odense, Heden 7-9, Indgang 20, DK – 5000 Odense C, Denmark. Tel: +45-26183619; Fax: +45-65415064; E-mail: karin.bruun.plesner@rsyd.dk.

Funding sources: KBP received funding for the study from the Danish Rheumatism Association (R155-A4821-B1326).

Conflicts of interest: The authors have no conflicts of interest to declare.

Abstract

Objective. This study explores dose–response relationships when treating fibromyalgia with low-dose naltrexone.**Design**. A single-blinded clinical trial was carried out using the "up-and-down" method.**Subjects**. Subjects included women with a diagnosis of fibromyalgia aged 18–60 years who had been referred to treatment at a public pain clinic at a Danish university hospital.**Methods**. The test doses were in the range 0.75–6 mg, and the dosing interval was 0.75 mg. The method was sequential and allowed predicting the dose effective in 50% (ED50) and 95% (ED95) of the subjects when the dose had shifted direction 10 times, and six pairs of "up-and-down" data were available.**Results**. A total of 27 subjects were included in the study; two subjects were withdrawn. After inclusion of 25 evaluable subjects, the dose estimates were calculated as 3.88 mg for ED50 and 5.40 mg for ED95. As a secondary outcome, the effects on 10 common fibromyalgia symptoms were evaluated. A high interindividual variation was observed both in the symptom presentation at baseline and in which symptoms were reduced by low-dose naltrexone.**Conclusions**. This study is the first to explore dose–response relationships in the treatment of fibromyalgia with low-dose naltrexone. Future placebo-controlled randomized clinical trials are needed, and according to our findings, 4.5 mg, which has previously been used, seems to be a relevant test dose. We recommend that future studies include additional nonpain fibromyalgia symptoms as outcome measures.

Key Words: Fibromyalgia; Low-Dose Naltrexone; LDN; Dose–Response

Introduction

Fibromyalgia (FM) is a chronic disorder characterized by generalized pain and tenderness, accompanied by a range of symptoms such as fatigue, insomnia, cognitive disturbances, increased sensitivity to other sensory inputs, anxiety, and depression [1]. FM is a common disorder affecting ~2% of the population [2]. Accumulating evidence indicates that FM patients have changes in their pain regulatory system, with facilitated excitation of nociceptors and reduced central inhibition, leading to a widespread hyperalgesic state [3]. More recently, disturbances in the immune system with signs of an inflammatory state in the central nervous system (CNS) have been found [3–6]. FM patients have been shown to have increased levels of enkephalins in the cerebrospinal fluid (CSF) [7], and decreased mu receptor availability has been demonstrated [8], suggesting that these patients have a dysfunction of the endogenous analgesic

mechanisms. No specific treatment of FM is available, and traditional pharmacological therapies aim at either reducing the release of facilitatory neurotransmitters (e.g., gabapentinoids) or blocking the reuptake of both serotonin and norepinephrine [1].

Low-dose naltrexone (LDN) is used widely as off-label treatment for various conditions [9,10], but the evidence for its use in FM is sparse [10]; to our knowledge, only three small clinical trials have been published [11–13].

Naltrexone is primarily known for its antagonistic effect on the opioid receptor [14], but it is also thought to blunt dopaminergic transmission in mesolimbic pathways [15], thereby attenuating craving and reinforcing effects of substances of abuse. Naltrexone is marketed as an additional therapy for supporting abstinence in patients with opioid or alcohol use disorders [16]. It has been known for decades that naltrexone can have a paradoxical analgesic effect when used in low doses [17]. The proposed mechanisms of action for LDN on FM are 1) improvement of opioid signaling and 2) an anti-inflammatory effect. Evidence, primarily from animal studies, has shown that LDN increases levels of both endorphin and met-enkephalin (opioid growth factor [OGF]) and increases expression of opioid receptors and OGF receptors [18,19]. Moreover, LDN binds to Toll-like receptor 4 (TLR4) on glial cells, where it exhibits antagonist properties, leading to reduction of pro-inflammatory cytokines such as interleukin-6 and tumor necrosis factor–alpha [13,20,21].

In the previous studies investigating the effect of LDN on FM, a dose of 4.5 mg was used. However, thus far, no dose–response studies have been carried out. The aim of this study was to explore dose–response relationships when treating FM patients with LDN, and thereby to estimate the dose effective in 50% of subjects (ED50) and the dose effective in 95% of subjects (ED95). A secondary aim was to investigate which FM symptoms were most commonly reduced by LDN.

Methods

Procedure

A single-center study was conducted at a public university hospital pain center in Southern Denmark. Approval was obtained from the Danish Data Protection Agency, the Ethical Committee of Southern Denmark (S-20160121), and the Danish Health and Medicines Authority. The study was registered with the European Union Drug Regulation Authorities Clinical Trials Database (EudraCT-nr: 2016–002081-31) and was monitored by the Good Clinical Practice (GCP) unit at Odense University Hospital. Informed consent was obtained from all subjects entering the study. Enrollment began in June 2017, and the study was completed in September 2018. The study medication was produced at Glostrup Pharmacy, and the tablets were blinded in similar white gelatin capsules and labeled with blinding codes. The medicine was administered orally once daily in the evening.

Entry Criteria

Subjects eligible for the study were Caucasian women aged 18–60 years who had a diagnosis of FM. The subjects were required to understand and write Danish, and all fertile women were requested to use secure anticonception or to be sexually abstinent for the three weeks before entering and the one week after concluding the trial. Exclusion criteria included known allergy toward naltrexone hydrochloride, any known inflammatory rheumatic disease, any known neurological disease, other significant localized pain conditions, psychiatric disease, suicidal ideation, suicidal attempt during the past five years, pregnancy or breastfeeding, use of opioids within the eight weeks before entering the trial, and abuse of alcohol or other substances.

All subjects included in the study were referred to treatment at the pain center and were identified at the end of the treatment course as patients who had not benefitted sufficiently from traditional pharmacological treatment. Because male patients with a diagnosis of FM are rarely referred to the pain center, it was decided to include only female subjects. Subjects eligible were examined by a pain specialist (KBP) at a screening interview, and all included subjects fulfilled both the classification criteria approved by the American College of Rheumatology (ACR) in 1990 [22] and the diagnostic criteria approved by the ACR in 2011 [23]. A pregnancy test was performed in all fertile subjects at inclusion.

Subjects were allowed to continue their usual medication during the trial, but the treatment had to be stable during the trial. Any change in usual medication or initiation of new medicine that could influence pain, sleep, or fatigue during the trial would lead to exclusion from the trial.

The subjects did not receive any compensation for participation in the trial.

Study Design

The study was designed as a prospective dose–response study using the "up-and-down" method [24]. The method is sequential and allows predicting the ED50 and ED95, when the dose has changed direction 10 times and six pairs of "up-and-down" data are available. These ED50 and ED95 serve as a range for a reasonable dose. The method usually requires 20–30 patients [24]. In this study, the dosing interval was 0.75 mg, and the chosen test doses were 0.75 mg, 1.5 mg, 2.25 mg, 3 mg, 3.75 mg, 4.5 mg, 5.25 mg, and 6 mg. All subjects received active treatment but were blinded to which dose they received. The primary investigator (KBP) was not blinded to which dose the subjects received.

The subjects had three visits during the trial, and all visits were performed by the primary investigator. At the first visit (day 1) baseline data were collected and medicine was dispensed in separate containers for the first two weeks of treatment (14 capsules) and for the third week of treatment (seven capsules). Visit 2 was a telephone visit (day 15), and visit 3 (day 29) was a follow-up meeting one week after the end of treatment. At visits 2 and 3, the subjects were interviewed about adverse events and compliance with the treatment. At visit 3, the empty medicine containers were returned, and any noningested capsules were counted. A subject was considered compliant if no more than two capsules were returned after the first two weeks of treatment and if no more than one capsule was returned after the third week of treatment. At all visits, the shared electronic medical record was updated to ensure that concomitant medication was stable. Electronic questionnaires were sent out by e-mail after two and three weeks of treatment. Because of the sequential method, subjects who were withdrawn or dropped out during the trial were not evaluated but were replaced.

The first subject was randomized to treatment with 2.25 mg, 3 mg, 3.75 mg, or 4.5 mg using the envelope method. Evaluation of response was made after two weeks of treatment based on the electronic questionnaires. The evaluation was made before visit 2 to reduce bias. The primary end point was subjective evaluation of improvement of overall FM symptoms using a seven-point transition scale—the Patient Global Impression of Improvement Scale (PGI-I). A subject was considered a responder if she scored 1–3 on the PGI-I scale, which ranges from 1 (very much improved) to 7 (very much worse), with 4 being no change. An additional primary end point was chosen to be a >30% reduction in pain from baseline, measured as average pain during the last three days on a 0–100 visual analog scale (VAS).

The primary end point was evaluated after two weeks, but the treatment was given for a total of three weeks to evaluate if significant changes in effect occurred between two and three weeks of treatment. If it turned out in the beginning of the trial that significant improvement occurred between two and three weeks of treatment, the protocol of the study allowed for changing the time for evaluation of the primary end point to three weeks instead of two weeks.

Secondary Outcomes

To evaluate the effect of LDN on other FM key symptoms, the 10 items from the symptom part of the Fibromyalgia Impact Questionnaire Revised (FIQR) [25] were chosen as secondary outcome measures. Each item evaluates the average severity of the symptom during the last seven days on a 0–10 numeric rating scale (NRS). The composite value of the 10 items from the symptom part of the FIQR (sFIQR) was calculated as a rough measure of symptom burden. The Insomnia Severity Index (ISI) questionnaire was used to evaluate changes in sleep quality [26].

Statistical Analyses

The primary analysis was made per protocol. Given that the "up-and-down" method [24] does not allow for imputation or related methods, subjects who were withdrawn or dropped out or who did not complete the treatment as intended were replaced. To account for the implicit assumption of increasing effect with increasing dose, the study applies isotonic regression, as suggested by Pace and Stylianou [24], which provides estimates of ED50 and ED95 with smaller bias than the traditional Dixon and Mood (1948) estimators [27]. The analysis was performed in STATA 15.0 using the IRAX module developed by van Putten and Royston [28]. As an extension of the isotonic regression, ED50 and ED95, together with their 95% confidence limits, were thus not calculated from data but predicted by the pooled-adjacent-violators algorithm (PAVA), as suggested by Pace and Stylianou [24]. Descriptive statistics (numbers, percentages, means, standard deviations, and interquartile ranges) were used to evaluate demographic data and secondary outcome measures.

Results

Patient Disposition

A total of 28 women were screened for the study, and 27 met the entry criteria. Two subjects were withdrawn from the trial, as they did not complete the treatment as intended. Both subjects were replaced. After inclusion of 25 evaluable subjects, the dose had changed direction 10 times, and six pairs of "up-and-down" data were obtained. Baseline data for all evaluable subjects and data on concomitant pain medication are presented in Table 1.

Primary Outcome

Of the 25 subjects analyzed, 11 were classified as responders, reporting either a minimum 30% decrease in pain from baseline or having a PGI-I score of 1–3 after two weeks of treatment. The up-and-down curve is shown in Figure 1, and the observed and fitted doses with 95% confidence intervals are shown in Table 2. In Figure 1, each circle is a patient, and the doses are shown in order from left to right. The first three patients did not respond, but the one given a dose of slightly over 5 did; hence the fifth patient was tried on a lower dose. That one failed, so the dose was increased to >5. Then patients 6, 7, 8, and 9 all succeeded, on increasingly lower doses, until number 10 failed, etc. The response rate (equal to the patient sequence number divided by the final sample size) is shown on the horizontal axis. The 95% confidence interval (CI) of Figure 1 and Table 2 is an estimated pointwise, partition-wise confidence interval

Table 1. Raw scores for demographics, clinical pain profile, level of the 10 items from the symptom part of the FIQR, symptom burden measured as composite value of the 10 items of the symptom part of the FIQR, degree of sleep disturbance measured by the ISI, and analgesic use at baseline

Demographic and Clinical Features at Baseline (N = 25)	Mean±SD (Range)	Interquartile Percentiles (25, 50, 75)
Age, y	47.0±9.4 (27–59)	(41, 49, 55)
BMI, kg/m^2	29.4±6.4 (18–42)	(24, 28, 32)
Pain duration, y	13.6±11.1 (1–44)	(7, 10, 20)
Average pain intensity at baseline (NRS: 0–100)	67.2±17.0 (14–94)	(59, 66, 82)
FIQR items, average value during last 7 d (0–10 NRS)		
Pain	7.0±1.4 (3–9)	(6–8)
Energy	7.4±1.7 (3–10)	(6, 8, 9)
Stiffness	6.7±1.9 (3–9)	(5, 7, 8)
Waking unrefreshed	7.9±2.1 (2–10)	(7, 8, 10)
Depressed	4.0±2.7 (0–9)	(2, 4, 6)
Concentration/memory	6.7±2.0 (1–9)	(6, 7, 8)
Anxiety	2.4±2.8 (0–10)	(0, 1, 5)
Tenderness to touch	7.6±2.2 (3–10)	(6, 8, 9)
Imbalance	3.9±2.8 (0–8)	(2, 4, 7)
Sensitivity to sensory inputs	6.1±2.8 (0–10)	(4, 7, 8)
Composite value for symptom part of FIQR at baseline (0–100)	59.6±12.0 (32–85)	(50, 60, 68)
ISI score at baseline (0–28)	17.6±5.0 (1–25)	(16, 18, 20)
Concomitant pain medication, No. using medication (%)		
Paracetamol users	16 (64)	
NSAID users	8 (32)	
TCA users	1 (4)	
Anticonvulsive users	2 (8)	
Baclofen users	11 (44)	
Tizanidine users	4 (16)	
SNRI users	6 (24)	
SSRI users	2 (8)	

BMI = body mass index; FIQR = Fibromyalgia Impact Questionnaire Revised; ISI = Insomnia Severity Index; NRS = numeric rating scale; NSAID = nonsteroidal anti-inflammatory drugs; SNRI = serotonin noradrenaline reuptake inhibitor; SSRI = selective serotonin reuptake inhibitor; TCA = tricyclic antidepressant.

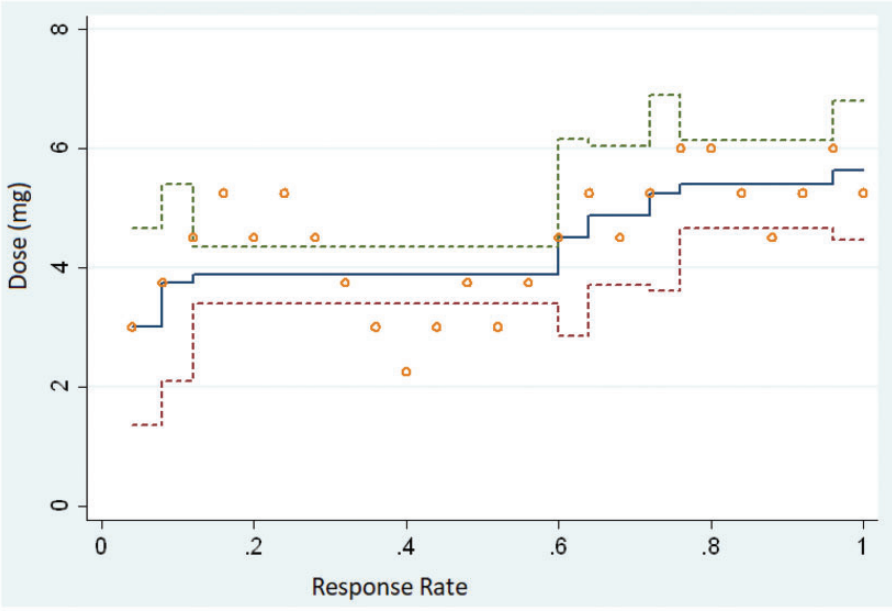


Figure 1. The up-and-down curve, showing isotonic regression of dose (mg) vs response rate with 95% confidence limits.

for the fitted step function, which IRAX fits by making auxiliary regression on the partition groups; see van Putten and Royston for details [28]. The PAVA estimates of μ_{50} and μ_{95} with 95% confidence limits were obtained as $\mu_{50} = 3.88$ (95% CI = 3.39 to 4.35) and $\mu_{95} = 5.40$ (95% CI = 4.66 to 6.13).

Of the 11 responders, one subject fulfilled only the pain reduction criteria, four subjects fulfilled both

Table 2. Observed and fitted doses (mg) with lower and upper 95% confidence limits

ID	Effect	Response Rate	Dose Observed	Dose Fitted	Lower 95% CI	Upper 95% CI
1	N	0.04	3.000	3.000	1.348	4.652
2	N	0.08	3.750	3.750	2.098	5.402
3	N	0.12	4.500	3.875	3.398	4.352
4	Y	0.16	5.250	3.875	3.398	4.352
5	N	0.20	4.500	3.875	3.398	4.352
6	Y	0.24	5.250	3.875	3.398	4.352
7	Y	0.28	4.500	3.875	3.398	4.352
8	Y	0.32	3.750	3.875	3.398	4.352
9	Y	0.36	3.000	3.875	3.398	4.352
10	N	0.40	2.250	3.875	3.398	4.352
11	N	0.44	3.000	3.875	3.398	4.352
12	Y	0.48	3.750	3.875	3.398	4.352
13	N	0.52	3.000	3.875	3.398	4.352
14	N	0.56	3.750	3.875	3.398	4.352
15	N	0.60	4.500	4.500	2.848	6.152
16	Y	0.64	5.250	4.875	3.707	6.043
17	N	0.68	4.500	4.875	3.707	6.043
18	N	0.72	5.250	5.250	3.598	6.902
19	N	0.76	6.000	5.400	4.661	6.139
20	Y	0.80	6.000	5.400	4.661	6.139
21	Y	0.84	5.250	5.400	4.661	6.139
22	N	0.88	4.500	5.400	4.661	6.139
23	N	0.92	5.250	5.400	4.661	6.139
24	Y	0.96	6.000	5.625	4.457	6.793
25		1.00	5.250	5.625	4.457	6.793

CI = confidence interval.

Table 3. Responder classification

Responders (N = 11)	30% Change in Pain After 2 Weeks (0–100 NRS;Yes/No)	Patient Global Impression of ChangeAfter 2 Weeks (1–7)	30% Change in Pain After 3 Weeks (0–100 NRS;Yes/No)	Patient Global Impression of Change After 3 Weeks (1–7)
1ID 4	Yes	4	Missing value	Missing value
2ID 6	No	3	No	3
3ID 7	No	1	No	1
4ID 8	Yes	2	Yes	2
5ID 9	Yes	2	Yes	2
6ID 12	Yes	3	Yes	3
7ID 16	No	3	No	3
8ID 20	Yes	2	No	2
9ID 21	No	2	No	3
10ID 24	No	3	No	3
11ID 25	No	2	No	3

This table shows how the responders were classified. Primary outcomes were Patient Global Impression of Improvement Score, ranging from 1 (very much better) to 7 (very much worse) and average daily pain during the past three days on a 0–100 NRS. A responder had to have either a 30% decrease in pain on a 0–100 NRS or a Patient Global Impression of Change Score between 1 and 3. All the patients who were classified as responders after two weeks also fulfilled the criteria after three weeks.

NRS = numeric rating scale.

criteria, and six subjects fulfilled only the PGI-I criteria (Table 3). After three weeks, 10 of the responders still fulfilled the criteria; for the last responder, data were lost due to problems with the electronic questionnaire. Three of the subjects who were classified as nonresponders after two weeks fulfilled the criteria for being responders after three weeks. All three subjects changed their PGI-I score from 4 (after two weeks) to 3 (after three weeks), and one of these patients also had a >30% decrease in average pain after three weeks compared with baseline (data not shown).

Secondary Outcomes

Data for the 11 responders are shown in Table 4. The two items from the FIQR with the highest mean change score after two weeks were self-perceived "tenderness," with a mean change of –2.3 (0–10 NRS) and five of 11 reporting a >30% improvement, and "waking unrefreshed," with a mean change of –2.3 and eight of 11 reporting a >30% improvement. For the item "pain," the mean change was only –1.4 (0–10 NRS), and only four of 11 reported a >30% improvement after two weeks. After three weeks, the improvement in pain was even less prominent. Most of

Table 4. Raw scores for symptom burden at baseline for the 11 responders, measured by mean score of each of the 10 items from the symptom part of the FIQR, the composite value of the 10 items of the FIQR, and mean changes in these scores after two and three weeks of treatment with LDN

Responders (N = 11)	Baseline (N = 11), Mean±SD (Range)	After 2 Weeks (N = 11), Mean Change±SD	After 3 Weeks (N = 10), Mean Change±SD	>30% Improvement After 2 Weeks, No./Total	>30% Improvement After 3 Weeks, No./Total
FIQR items, average value during last 7 d (0–10 NRS)					
Pain	6.7±1.5 (3–9)	–1.4±1.3	–0.4±2.0	4/11	2/10
Energy	7.4±2.1 (3–10)	–1.5±2.0	–1.7±2.3	4/11	4/10
Stiffness	6.3±2.1 (3–9)	–0.6±2.5	–1.2±2.4	2/11	4/10
Waking unrefreshed	7.8±2.0 (3–10)	–2.3±2.1	–2.2±1.8	8/11	5/10
Depressed	3.7±3.1 (0–9)	–1.5±2.0	–1.2±2.1	6/11	4/10
Concentration/memory	6.3±1.9 (3–9)	–1.2±1.9	–1.2±2.4	4/11	4/10
Anxiety	2.0±2.5 (0–7)	–0.5±1.1	–1.3±2.5	3/11	4/10
Tenderness	7.2±2.1 (3–10)	–2.3±1.9	–3.5±1.8	5/11	8/10
Imbalance	3.8±2.9 (0–8)	0, 5±2.5	–0.8±1.8	1/11	4/10
Sensitivity to sensory inputs	5.5±2.8 (0–9)	–0.3±2.1	–0.2±2.6	2/11	3/10
Composite value of the 10 items from the symptom part of FIQR (0–100)	56.6±13.6	–10.8±9.3	–13.6±12.1	2/11	5/10
Insomnia Severity Score (0–28)	16.9±2.6	–3.3±4.6	–5.4±4.6	4/11	4/10

FIQR = Fibromyalgia Impact Questionnaire Revised; LDN = low-dose naltrexone; NRS = numeric rating scale.

the other items remained stable or improved further after three weeks. The item self-perceived "tenderness," especially, was further improved after three weeks, with a mean change of –3.5 (0–10 NRS) from baseline and eight of 11 reporting a >30% improvement.

All responders had a minimum improvement of 30% on at least one of the 10 FM symptoms, but most of the patients showed a minimum improvement of 30% on several symptoms. This is illustrated by the mean change in the composite value of the 10 items from the sFIQR, which showed a mean value of 56.6 ± 13.6 (0–100 NRS) at baseline, with a mean change of –10.8 ± 9.3 after two weeks and a mean change of –13.6 ± 12.1 after three weeks. After two weeks of treatment, the average number of domains with >30% improvement was 3.5, and after three weeks, the average number of domains improved was 4.2. Nonresponders (N = 14) had a mean change of 3.3 ± 9.9 in sFIQR score after two weeks and a mean change of 1.1 ± 10.2 on the sFIQR after three weeks (data not shown).

Adverse Events

No serious adverse events occurred during the trial, but side effects were common. Two subjects were withdrawn from the study because of noncompliance. Both the subjects had many side effects. One reported severe nausea, abdominal pain, and headache and scored 6 on the PGI-I scale after two weeks and did not want to continue the treatment for the last week; the other only ingested five capsules and reported fatigue, depression, headache, and abdominal pain to be the reason for withdrawal from the study. Both patients received doses on the low end of the dose range (3 mg and 3.75 mg). Adverse events were common but were generally graded mild and tolerable. Gastrointestinal symptoms were the most commonly reported side effects. The side effects reported after two weeks of treatment are listed in Table 5.

Discussion

In this study, we explored dose–response relationships when treating FM with LDN using the "up-and-down" method. We estimated the ED50 to be 3.88 mg and the ED95 to be 5.40 mg. Since its introduction in the 1980s [29], LDN has been used widely as off-label treatment for fibromyalgia and other chronic pain conditions [10]. Several case reports have been published reporting a pain-relieving effect of LDN [30–32]. The evidence is sparse, however, with few clinical trials. The doses used in the published case reports varied from 1 mg to 5 mg [10], but in all previous clinical trials testing the effect of LDN on FM, a dose of 4.5 mg has been used [12]. To our knowledge, no dose–response studies have ever been published. For future studies of the effect of LDN on FM, it is important to estimate a dose that is sufficiently effective for many patients and still not too high to give rise to intolerable side effects, which would lead to high dropout rates and poor quality of the studies. Based on our findings in this study, a dose of 4.5 mg seems to be a reasonable test dose in FM patients, as it lies in the range between our estimates of ED50 and ED95.

Given the sequential method, a relatively short treatment period was needed. Time-effect curves from previous studies [11,12] show that FM symptoms seem to improve during the first two weeks of treatment, with some further improvement during the following six to 10 weeks. Previous studies have shown that 57–60% of patients get >30% relief of pain when treated with LDN

Table 5. Reported side effects after two weeks of treatment

Reported Side Effects (27 Subjects)	Dose Received, 2.25 mg (1 Subject)	Dose Received, 3 mg (5 Subjects)	Dose Received, 3.75 mg (5 Subjects)	Dose Received, 4.5 mg (6 Subjects)	Dose Received, 5.25 mg (7 Subjects)	Dose Received, 6 mg (3 Subjects)
Abdominal ache	1	1	1	2	2	
Diarrhea		1		1	1	
Constipation					2	
Nausea		1	2			
Headache		2	1	1		
Vivid dreams			1			
Mood disturbance			1			
Increased pain					1	
Increased fatigue			1			
Dizziness	1					
Palpitations		1				
Increased appetite						1
Sleeping difficulty				1		

The subjects received different doses; the number of subjects reporting a given side effect is listed for each dose. Some subjects reported more than one side effect. Side effects were common, and only six of 27 subjects did not report any side effects.

for eight to 12 weeks [11,12]. We found that only a smaller percentage of the responders in this study reported a reduction >30% in pain after two weeks, and the average reduction of pain after three weeks was minimal. Instead we found that the FIQR item self-perceived "tenderness" was the symptom most reduced on average after both two and three weeks of treatment. The design of this study does not allow us to make any conclusions about secondary outcome measures, but it does give rise to some hypothetical considerations. The decrease in self-perceived tenderness could reflect an improvement of the hyperalgesic state seen in FM patients, and we hypothesize that this, together with improvement of energy and sleep, might lead to increased activity, which could blunt the expected improvement of pain. Previous studies exploring the effect of LDN on FM have been longer, and patients might find a new balance in daily activities over a longer period of time, which might explain that studies with longer duration find pain to be significantly reduced by LDN. The dose of LDN may also play a role, as previous studies have used 4.5 mg, and some of the responders in this study received doses below that.

It is well known that there is high interindividual variability in which symptoms FM patients report to be most severe [33]. From clinical experience, there is also interindividual variability in which symptoms are relieved by LDN. Many patients report improvement of sleep and/or fatigue instead of pain relief as the major benefit from the treatment. In this study, we also found a great variability in both symptoms at baseline and, as an expected consequence of this, interindividual variability in which symptoms were relieved by LDN. In line with the IMMPACT guidelines [34], we therefore recommend that future studies on the effect of LDN on FM include phenotyping of the subjects to account for possible variations in pathophysiology and to include outcome measures of key FM symptoms other than pain, as well as measures of changes in function.

Limitations

Our results might be biased by several factors. Given the sequential method, we had to evaluate the effect of the treatment after a relatively short period of time, and two weeks was chosen based on time–response curves from previous trials. This period of time might not have been sufficient to give a positive effect in all patients, as the effect of LDN is thought to be mediated through opioid receptor upregulation and attenuation of inflammatory pathways, which might require more time. We found that three patients reported positive effect after three weeks but not after two weeks. During the trial, we did not find these three cases sufficient to change the evaluation of response at three weeks, but if the evaluation had been made after three weeks, it might have given rise to a lower estimated ED50 and ED95. The test doses in the study were chosen based on our clinical experience that many FM patients benefit from lower doses than 4.5 mg, and as we did not have any experience using doses higher than 4.5 mg, we were reluctant to choose doses higher than 6 mg. During the study, it became clear that the subjects tolerated doses up to 6 mg well. If we had chosen to include test doses above 6 mg, we might have found a higher estimated ED50 and ED95. Another limitation to the trial is that some patients might have experienced a placebo effect after only two weeks of treatment. It should be noted, though, that all responders reported a positive effect after both two and three weeks of treatment, and all responders had a minimum 30% decrease of several common FM symptoms, with the average number of domains improved after two and three weeks being 3.5 and 4.2, respectively. The subjects in this study were drawn from a university hospital, and the subjects had typically failed to benefit from traditional therapies. These subjects belong to the severe end of the spectrum of FM and might benefit less from treatment with LDN than FM patients with a milder level of the disease. The study was only single-blinded. Double-blinding would

have been preferable but was not feasible due to limited funding. To minimize bias, subjects completed their questionnaires electronically before the visit, thereby minimizing investigator influence on response. This study was designed as a dose–response study, and evaluation of effect parameters must be made with caution.

Conclusions

In this dose–response study investigating the effect of LDN on FM using the "up-and-down" method, we have estimated the ED50 to be 3.88 mg and the ED95 to be 5.40 mg. Larger randomized controlled trials comparing LDN with placebo are needed in the future, allowing for large-sample traditional statistics like chi-square-based tests for effects. Given that the "up-and-down" sample method does not support sample size or power calculations for such a trial, a pilot trial might conveniently precede. Choosing a relevant test dose is crucial, as a test dose that is too low might lead to low response rates and a test dose that is too high might lead to high dropout rates because of side effects. Based on our current findings, we conclude that 4.5 mg, which has been used in previous trials, seems to be a good choice, as it lies between our estimates for ED50 and ED95.

We suggest that future clinical trials exploring the effect of LDN on FM incorporate measures of "tenderness," valid and reliable measures of hyperalgesia, and measures of physical functioning to explore the hypothesis that LDN primarily influences hyperalgesia, fatigue, and sleep and secondarily influences pain.

References

1. Clauw DJ. Fibromyalgia: A clinical review. JAMA 2014;311(15):1547–55.
2. Wolfe F, Walitt B, Perrot S, Rasker JJ, Hauser W. Fibromyalgia diagnosis and biased assessment: Sex, prevalence and bias. PLoS One 2018;13(9):e0203755.
3. Sluka KA, Clauw DJ. Neurobiology of fibromyalgia and chronic widespread pain. Neuroscience 2016;338:114–29.
4. Backryd E, Tanum L, Lind AL, Larsson A, Gordh T. Evidence of both systemic inflammation and neuroinflammation in fibromyalgia patients, as assessed by a multiplex protein panel applied to the cerebrospinal fluid and to plasma. J Pain Res 2017;10:515–25.
5. Uceyler N, Hauser W, Sommer C. Systematic review with meta-analysis: Cytokines in fibromyalgia syndrome. BMC Musculoskelet Disord 2011;12:245–39.
6. Rodriguez-Pinto I, Agmon-Levin N, Howard A, Shoenfeld Y. Fibromyalgia and cytokines. Immunol Lett 2014;161(2):200–3.
7. Baraniuk JN, Whalen G, Cunningham J, Clauw DJ. Cerebrospinal fluid levels of opioid peptides in fibromyalgia and chronic low back pain. BMC Musculoskelet Disord 2004;5(1):48–55.
8. Gracely RH, Ambrose KR. Neuroimaging of fibromyalgia. Best Pract Res Clin Rheumatol 2011;25(2):271–84.
9. Raknes G, Smabrekke L. A sudden and unprecedented increase in low dose naltrexone (LDN) prescribing in Norway. Patient and prescriber characteristics, and dispense patterns. A drug utilization cohort study. Pharmacoepidemiol Drug Saf 2017;26(2):136–42.
10. Toljan K, Vrooman B. Low-dose naltrexone (LDN)—review of therapeutic utilization. Med Sci (Basel) 2018;6(4):82–100.
11. Younger J, Mackey S. Fibromyalgia symptoms are reduced by low-dose naltrexone: A pilot study. Pain Med 2009;10(4):663–72.
12. 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(2):529–38.
13. Younger J, Parkitny L, McLain D. The use of low-dose naltrexone (LDN) as a novel anti-inflammatory treatment for chronic pain. Clin Rheumatol 2014;33(4):451–9.
14. Verebey K. The clinical pharmacology of naltrexone: Pharmacology and pharmacodynamics. NIDA Res Monogr 1981;28:147–58.
15. Valenta JP, Job MO, Mangieri RA, Schier CJ, Howard EC, Gonzales RA. μ-opioid receptors in the stimulation of mesolimbic dopamine activity by ethanol and morphine in Long-Evans rats: A delayed effect of ethanol. Psychopharmacology (Berl) 2013;228(3):389–400.
16. Yoon G, Kim SW, Thuras P, Westermeyer J. Safety, tolerability, and feasibility of high-dose naltrexone in alcohol dependence: An open-label study. Hum Psychopharmacol 2011;26(2):125–32.
17. Gillman MA, Lichtigfeld FJ. A pharmacological overview of opioid mechanisms mediating analgesia and hyperalgesia. Neurol Res 1985;7(3):106–19.
18. Tempel A, Kessler JA, Zukin RS. Chronic naltrexone treatment increases expression of preproenkephalin and preprotachykinin mRNA in discrete brain regions. J Neurosci 1990;10(3):741–7.
19. Tempel A, Gardner EL, Zukin RS. Neurochemical and functional correlates of naltrexone-induced opiate receptor up-regulation. J Pharmacol Exp Ther 1985;232(2):439–44.
20. Wang X, Zhang Y, Peng Y, et al. Pharmacological characterization of the opioid inactive isomers (+)-naltrexone and (+)-naloxone as antagonists of toll-like receptor 4. Br J Pharmacol 2016;173(5):856–69.
21. Cant R, Dalgleish AG, Allen RL. Naltrexone inhibits IL-6 and TNFalpha production in human immune cell subsets following stimulation with ligands for intracellular toll-like receptors. Front Immunol 2017;8:809.
22. 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(2):160–72.
23. Wolfe F, Clauw DJ, Fitzcharles MA, et al. Fibromyalgia criteria and severity scales for clinical and epidemiological studies: A modification of the ACR Preliminary Diagnostic Criteria for Fibromyalgia. J Rheumatol 2011;38(6):1113–22.
24. Pace NL, Stylianou MP. Advances in and limitations of up-and-down methodology: A precis of clinical use, study design, and dose estimation in anesthesia research. Anesthesiology 2007;107(1):144–52.
25. 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(4):415–29.
26. Bastien CH, Vallieres A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001;2(4):297–307.

27. Stylianou M, Flournoy N. Dose finding using the biased coin up-and-down design and isotonic regression. Biometrics 2002;58 (1):171–7.
28. Van Putten W, Royston P. IRAX: Stata Module to Perform Isometric Regression Analysis. Boston: Boston College Department of Economics; 2017.
29. Bihari B. Low-dose naltrexone for normalizing immune system function. Altern Ther Health Med 2013;19(2): 56–65.
30. Ghai B, Bansal D, Hota D, Shah CS. Off-label, low-dose naltrexone for refractory chronic low back pain. Pain Med 2014;15(5): 883–4.
31. Hota D, Srinivasan A, Dutta P, Bhansali A, Chakrabarti A. Off-label, low-dose naltrexone for refractory painful diabetic neuropathy. Pain Med 2016;17(4):790–1.
32. Sturn KM, Collin M. Low-dose naltrexone: A new therapy option for complex regional pain syndrome type I patients. Int J Pharm Compd 2016;20(3):197–201.
33. Toussaint L, Vincent A, McAllister SJ, Whipple M. Intra- and inter-patient symptom variability in fibromyalgia: Results of a 90-day assessment. Musculoskeletal Care 2015;13(2):93–100.
34. Edwards RR, Dworkin RH, Turk DC, et al. Patient phenotyping in clinical trials of chronic pain treatments: IMMPACT recommendations. Pain 2016;157(9):1851–71.

Bruun et al. Trials (2021) 22:804
https://doi.org/10.1186/s13063-021-05776-7

Trials

STUDY PROTOCOL Open Access

Low-dose naltrexone for the treatment of fibromyalgia: protocol for a double-blind, randomized, placebo-controlled trial

Karin Due Bruun[1,2,3*], Kirstine Amris[3,4], Henrik Bjarke Vaegter[1,2], Morten Rune Blichfeldt-Eckhardt[1,2], Anders Holsgaard-Larsen[2,5], Robin Christensen[2,3] and Palle Toft[2,6]

Abstract

Background: Low-dose naltrexone (LDN) is used widely as an off-label treatment for pain despite limited evidence for its effectiveness. A few small trials with a high risk of bias have investigated the effect of LDN on pain associated with fibromyalgia in women, but larger and more methodologically robust studies are needed. The primary aim of this randomized controlled trial is to investigate if 12 weeks of LDN treatment is superior to placebo in reducing the average pain intensity during the last 7 days in women with fibromyalgia.

Methods: A single-center, permuted block randomized, double-blind, placebo-controlled, parallel-group trial will be performed in Denmark. Randomization comprises 100 women aged 18–64 years diagnosed with fibromyalgia who will be treated with either LDN or placebo for 12 weeks including a 4-week titration phase. The primary outcome is change in average pain intensity (during the last 7 days) from baseline to 12 weeks. Secondary outcomes are other fibromyalgia-related symptoms, i.e., tenderness, fatigue, sleep disturbance, stiffness, memory problems, depression, anxiety and measures of global assessment, physical function, impact of fibromyalgia, pain distribution, and health-related quality of life. Intention-to-treat analysis will be performed, and the number of responders with a more than 15%, 30%, and 50% improvement of pain after 12 weeks will be calculated for the LDN and placebo groups. Exploratory outcomes include measures of pain sensitivity, muscle performance, and biomarkers.

Discussion: This study will contribute with high-level evidence on the efficacy of low-dose naltrexone for the treatment of pain in women with fibromyalgia. Secondary outcomes include both disease-specific and generic components investigating whether LDN influences other symptoms than pain. Explorative outcomes are included to provide greater insight into the mechanism of action of LDN and possibly a better understanding of the underlying pathology in fibromyalgia.

Trial registration: EudraCT 2019-000702-30. Registered on 12 July 2019. ClinicalTrials.gov NCT04270877. Registered on 17 February 2020

Keywords: Fibromyalgia, Pain, Low dose naltrexone, LDN, RCT

* Correspondence: karin.due.bruun@rsyd.dk
[1]Pain Research Group, Pain Center, Odense University Hospital, Heden 7-9, Indgang 20, DK – 5000 Odense C, Denmark
[2]Department of Clinical Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark
Full list of author information is available at the end of the article

Administrative information

Note: the numbers in curly brackets in this protocol refer to SPIRIT checklist item numbers. The order of the items has been modified to group similar items (see http://www.equator-network.org/reporting-guidelines/spirit-2013-statement-defining-standard-protocol-items-for-clinical-trials/).

Title {1}	Low dose naltrexone for the treatment of fibromyalgia: Protocol for a double-blind, randomized, placebo-controlled trial
Trial registration {2a and 2b}.	EudraCT-number.: 2019-000702-30 ClinicalTrials.gov Identifier: NCT04270877
Protocol version {3}	Version 5.0 (25.09.2020). Amendment: Version 5.1 (27.07.2021)
Funding {4}	The Danish Rheumatism Association has granted 175.000 DKK for operating expenses. Odense University Hospital's PhD fund for operation expenses has granted 100.000 DKK towards operating expenses and the study nurse's salary. Aase and Ejnar Danielsens fund has granted 100.000 DKK towards operating expenses and the study nurse's salary. The Parker Institute at Bispebjerg and Frederiksberg Hospital is supported by a core grant from the Oak Foundation (OCAY-18-774-OFIL). The work of the Primary Investigator (KDB) is covered by a PhD salary.
Author details {5a}	Karin Due Bruun*[1,2,3] karin.due.bruun@rsyd.dk Kirstine Amris[3,4] kirstine.amris@regionh.dk Henrik Bjarke Vaegter[1,2] hbv@rsyd.dk Morten Rune Blichfeldt-Eckhardt[1,2] mr.be@rsyd.dk Anders Holsgaard-Larsen[2,4] AHLarsen@health.sdu.dk Robin Christensen[2,3] robin.christensen@regionh.dk Palle Toft[2,6] palle.toft@rsyd.dk [1] Pain Research Group, Pain Center, Odense University Hospital, Odense, Denmark [2] Department of Clinical Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark [3] The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark 4 Department of Rheumatology, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark [5] Department of Orthopedics and Traumatology, Odense University Hospital, Odense, Denmark [6] Department of Anesthesiology and Intensive Care, Odense University Hospital, Odense, Denmark
Name and contact information for the trial sponsor {5b}	Palle Toft, Professor, MD, PhD, Dr. Med. Department of Anesthesiology and Intensive Care Odense University Hospital Sdr. Boulevard 29 5000 Odense C Email: palle.toft@rsyd.dk
Role of sponsor {5c}	The study is investigator initiated.

Introduction

Background and rationale {6a}

Low-dose naltrexone (LDN) has been used as an off-label treatment for pain and inflammation in multiple sclerosis, Crohn's disease, and fibromyalgia (FM) for several years [1]. Naltrexone (NLX) is marketed as an additional therapy for the prevention of relapse in patients with previous abuse of opioids or alcohol [2]. While it is primarily known as an opioid receptor antagonist [3], NLX also attenuates dopaminergic transmission in mesolimbic pathways, thereby reducing cravings after substance abuse [4]. NLX has a similar biochemical structure to Naloxone but a higher oral bioavailability and a longer half-life [5], and it is well known that NLX can have a paradoxical analgesic effect when used in low doses of 1–6 mg [6].

The proposed mechanisms of action of LDN on pain are 1) opioid antagonism, which leads to a feedback-mediated increased expression of opioid receptors in the central nervous system (CNS) [7, 8] with a possible improvement of the endorphin system and 2) an anti-inflammatory effect, mediated through inhibition of Toll-like receptor 4 (TLR4) on astrocytes and microglia cells, thereby possibly inhibiting the pro-inflammatory cytokine cascade thought to be involved in the development and maintenance of chronic pain [9, 10].

The evidence for an analgesic effect of LDN is sparse, however. Several case reports exist [11–13], but only three small clinical trials have been published. The first trial was a single-blind pilot study with participation of 10 women with FM [14]. The subjects received placebo for 2 weeks, followed by an 8-week treatment with LDN 4.5 mg. Quantitative sensory testing showed improved pressure pain and heat pain thresholds during treatment with LDN compared to placebo. The same research team conducted a double-blind, placebo-controlled, randomized (cross-over) trial (RCT) [15], where 31 women with FM were randomized to receive either 4-week treatment with placebo followed by 12-week treatment with LDN 4.5 mg or 12-week treatment with LDN 4.5 mg followed by 4-week treatment with placebo. Both studies found LDN to be significantly better than placebo in reducing pain. The third and most recent study was a single-blind non-controlled pilot study with participation of 8 women with FM [16]. The participants were told they could receive placebo at any time during

the 8-week intervention, but all patients received active treatment (LDN 4.5 mg) throughout the trial. Significant reductions from baseline were seen in 17 out of 63 pro-inflammatory cytokines, supporting the hypothesis of an anti-inflammatory effect of LDN. The two pilot studies represent important pioneer work, but do not provide high-level evidence because of lack of power and single-blind or non-controlled study designs with a high risk of bias. In the cross-over trial, the method is more robust with both randomization and double-blinding. However, the study also has some weaknesses. Although showing promising results, it is unclear if the study was sufficiently powered and the decision to exclude a washout period between the interventions increases the risk of bias.

LDN has been shown to be a safe treatment [17] and a low-cost alternative to traditional therapies, but larger RCTs are needed to confirm its potential efficacy in reducing pain in patients diagnosed with fibromyalgia. Previous trials investigating the effect of LDN on pain have used one daily dose of 4.5 mg. However, higher doses might be more beneficial for some patients. Our study group previously conducted a dose-response study testing doses in the range of 0.75–6 mg [18]. We found the effective dose in 50% (ED50) to be 3.88 and the effective dose in 95% (ED95) to be 5.40 mg. We concluded that 4.5 mg would be a relevant test dose as it lies in the range between ED50 and ED95. However, doses closer to ED95 would be expected to be even more efficacious. As we found no problems with tolerability using doses in the range from 4.5 to 6 mg, we decided to test 6 mg against placebo in this RCT.

Objectives {7}

The primary objective is to investigate if 12 weeks' treatment with 6 mg LDN is superior to placebo in reducing the average pain intensity (during the last 7 days) in women with fibromyalgia. Secondary objectives include evaluating the clinical effect on 21 secondary outcomes covering core symptoms, daily functioning, impact of FM, quality of life, global impression of change, and responder indices. Finally, we will explore effects on pressure pain thresholds, temporal summation of pain, conditioned pain modulation, physical fitness, muscle exhaustion, and blood levels of pro-inflammatory cytokines.

Trial design {8}

The study is designed as a single-center, permuted block randomized, double-blind, placebo-controlled, parallel-group trial. Randomization comprises a parallel randomized (1:1) allocation of 100 women aged 18–64 years diagnosed with fibromyalgia, treated with either LDN or placebo for 12 weeks including a 4-week titration phase (from baseline to week 4).

Methods: participants, interventions, and outcomes

Study setting {9}

The study is a single-center study that is conducted at a public university hospital in Southern Denmark (SMERTECENTER SYD, Heden 7-9, 5000 Odense C). The setting is a tertiary pain rehabilitation center.

Eligibility criteria {10}

Inclusion criteria:

- Women aged 18–64 years
- Can understand and write Danish
- Fulfill the American College of Rheumatology 1990 criteria for FM [19]
- A minimum score of 4 for self-reported average pain during the last 7 days on a 0–10 numeric rating scale (NRS) at baseline
- Women of child-bearing age must use safe contraception (spiral, birth control pills, contraceptive patch, contraceptive vaginal ring, or gestagen injections) for 3 weeks before and 1 week after the trial. If a participant's usual lifestyle includes sexual abstinence, contraception is not required, but the participant must give oral informed consent that they will remain sexually abstinent during the trial

Exclusion criteria:

- Known allergy to naltrexone hydrochloride
- Pregnancy or breastfeeding; a negative pregnancy test must be available at baseline for all women of fertile age
- Use of opioids or NSAIDs up to 4 weeks before inclusion in the trial
- Known abuse of alcohol or other substances
- Known inflammatory rheumatic disease
- Known demyelinating disease
- Known active cancer
- Liver dysfunction (alanine aminotransferase (ALAT) must not be elevated more than 2-fold over the highest reference level)
- Kidney dysfunction (glomerular filtration rate (GFR) must not be below 59 mL/min)
- Psychotic disease
- History of a suicide attempt
- Suicide ideation—evaluated using Patient Health Questionnaire—9 items (PHQ-9) [20]; item 9 must be answered "never"

Who will take informed consent? {26a}

Potential participants recruited from the pain center will receive written information about the trial from their nurse or physician. For potential participants recruited

via advertising, written information is sent by e-mail. All potential participants receive a telephone call from the primary investigator (PI) (author KDB), who gives oral information about the trial. It is emphasized that participation is voluntary and that consent can be withdrawn at any time. A minimum of 24 h is given for reflection. The PI obtains the informed consent before inclusion.

Additional consent provisions for collection and use of participant data and biological specimens {26b}

Informed consent to use blood from the biobank to perform analyses for other research purposes is obtained from all participants.

Interventions

Explanation for the choice of comparators {6b}

No comparators are used other than the identically appearing placebo control.

Intervention description {11a}

After inclusion, the participants will be randomized using a computerized algorithm to receive either placebo or LDN for 12 weeks. The participants' dose will be titrated up to 6 mg following a dose-escalation scheme: an initial dosage of 1.5 mg daily, escalated every seventh day by 1.5 mg up to 6 mg at week 4. Dose escalation will be based on safety and tolerability, and if dose escalation is not feasible, delayed increments are allowed. For the surveillance of harms, both active and passive methods will be used. The participants will be encouraged to report adverse events spontaneously and will be asked about the occurrence of specific common side effects by administering a questionnaire. For the graduation of the severity of harms, the Common Terminology Criteria for Adverse Events (CTCAE) version 5.0 will be used. If the participant reports harms categorized as grade 2 or higher, they will be advised by the primary investigator to lower the dose. If harms are categorized as grade 1, the decision about dosing will be made individually in agreement between the primary investigator and the participant. After the end of week 4, the dose will be fixed for the rest of the trial, as the highest dose tolerated at this time point. The trial medicine is taken once daily in the evening, between 7 pm and 11 pm.

Criteria for discontinuing or modifying allocated interventions {11b}

The participants will be maintained at 6 mg (or the highest tolerated dose level established after the end of week 4) for the last 8 weeks of the treatment period. It is not allowed to increase the dose during the last 8 weeks. If problems with tolerability should arise during the last 8 weeks of treatment, it is allowed to lower the dose or discontinue treatment. Participants who alter the dose during the last 8 weeks of the trial will be considered not adherent to the protocol, but will be included in the intention-to-treat analysis.

Strategies to improve adherence to interventions {11c}

Participants will receive a daily short text message (SMS) reminding them to take their trial medication. At all visits, empty medicine cans are returned, and non-ingested tablets are counted.

Relevant concomitant care permitted or prohibited during the trial {11d}

The use of opioids, NSAIDs, and other drugs with an anti-inflammatory effect is prohibited during the trial. Participants can continue their usual care during the trial, but their pain medication has to be stable. The participants are not allowed to receive any new pain medication during the trial. Changes in concomitant medication are monitored at every visit via the patient's shared electronic medication record.

Provisions for post-trial care {30}

In the case of adverse events or adverse reactions, the PI will follow up on the participants until the symptoms have ceased or are stable. The participants are covered by the governmental patient insurance, which covers all patients in the Danish health care system.

Outcomes {12}

As previous studies have shown significant reductions in pain intensity in women with FM treated with LDN 4.5 mg for 8–12 weeks, we have chosen the primary outcome to be change in average pain intensity (during the last 7 days) from baseline to 12 weeks of intervention. The 21 secondary outcome measures were chosen among measures that could potentially support a clinical effectiveness claim as recommended by the Outcome Measures in Rheumatology Clinical Trials (OMERACT) guidelines [21]. All patient-reported outcomes will be collected at baseline and after 4, 8, and 12 weeks.

Primary outcome measure

Change from baseline to 12 weeks of treatment in average pain intensity during the last 7 days on an 11-point rating scale (ranging from 0 = "no pain" to 10 = "unbearable pain") using the first item from the symptom part of the Fibromyalgia Impact Questionnaire Revised (FIQR) [22].

Secondary outcome measures

The secondary outcomes include 21 supportive measures that will be collected, analyzed, and reported in the primary manuscript.

For the following secondary outcomes, the between-group change at baseline compared to 4, 8, and 12 weeks of treatment will be assessed:

1. Global assessment: assessed by Patient Global Impression of Change on a 1–7 Verbal Rating Scale
2. Impact of fibromyalgia: assessed by the FIQR total score [22]
3. Pain distribution: assessed by the Widespread Pain Index (WPI) from the 2016 diagnostic criteria for fibromyalgia [23]
4. Level of pain (assessment of pain intensity trajectory): assessed by the FIQR "level of pain" question
5. Level of tenderness: assessed subjectively by the FIQR "level of tenderness to touch" question and objectively by measurement of pressure pain threshold, using a handheld algometer. Algometry is performed only at baseline and after 12 weeks of treatment
6. Level of fatigue: assessed by the FIQR "level of energy" question
7. Level of sleep disturbance: assessed by the FIQR "quality of sleep" question
8. Level of depression: assessed by the FIQR "level of depression" question
9. Level of anxiety: assessed by the FIQR "level of anxiety" question
10. Level of cognition: assessed by the FIQR "level of memory problems" question
11. Level of stiffness: assessed by the FIQR "level of stiffness" question
12. Level of physical function: assessed by the physical function domain of FIQR
13. Health-related quality of life - mobility: assessed by the EQ-5D-5L mobility domain
14. Health-related quality of life - self-care: assessed by the EQ-5D-5L self-care domain
15. Health-related quality of life - usual activities: assessed by the EQ-5D-5L usual activities domain
16. Health-related quality of life - pain/discomfort: assessed by the EQ-5D-5L pain/discomfort domain
17. Health-related quality of life - anxiety/depression: assessed by the EQ-5D-5L anxiety/depression domain
18. Health-related quality of life - global: assessed by the EQ-5D Visual Analogue Scale (EQ-VAS)

Responder indices are calculated:

19. Number of responders with a more than 15% improvement of the primary outcome
20. Number of responders with a more than 30% improvement of the primary outcome
21. Number of responders with a more than 50% improvement of the primary outcome

Exploratory secondary outcomes (not to be reported in the primary manuscript)

The following exploratory outcomes will be investigated and reported in secondary publications. For the patient-reported outcome (variation in pain), the between-group change between baseline and after 8 and 12 weeks of treatment is measured. For all the protocol-specific procedures, the between-group change between baseline and after 12 weeks of treatment is measured.

- Variation in pain: assessed using a diary of daily average pain rated on an 11-point rating scale during 7 days before visits. The highest score minus the lowest score characterizes the variation in pain
- Muscle exhaustion: measured by an isometric muscle exhaustion test of the deltoid muscle
- Physical fitness: measured by the 30-s chair stand test
- Pain sensitivity: measured by computerized pressure cuff algometry (CPA)
- Inhibition of pain: measured by CPA using conditioned pain modulation (CPM)
- Augmentation of pain: measured by CPA using temporal summation of pain (TSP)

Blood for a biobank will be collected before baseline and immediately after 12 weeks of treatment for later analysis of pro- and anti-inflammatory cytokines. A separate protocol will be made to determine which cytokines will be investigated before the analyses are carried out.

Participant timeline {13}

The participant flow is shown in Fig. 1. A time schedule for enrolment, interventions, and assessments is presented in Table 1.

Sample size {14}

Using values from our previous dose-response study [18], we determined that self-reported pain on a 0–10 NRS at baseline had a mean of 6.7 in the target population, with a standard deviation (*SD*) of 1.5 NRS points. According to the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) guidelines [24], a minimal clinical important difference (MCID) is defined as a 15% decrease in pain [24], corresponding to a reduction of 1.0 NRS points in the present population. Using an MCID of 1.0 NRS, an *SD* of 1.5, a statistical power of at least 80%, and a statistical significance level of 0.05, a total of 74 patients are required, i.e., 37 patients in each group. Expecting some attrition

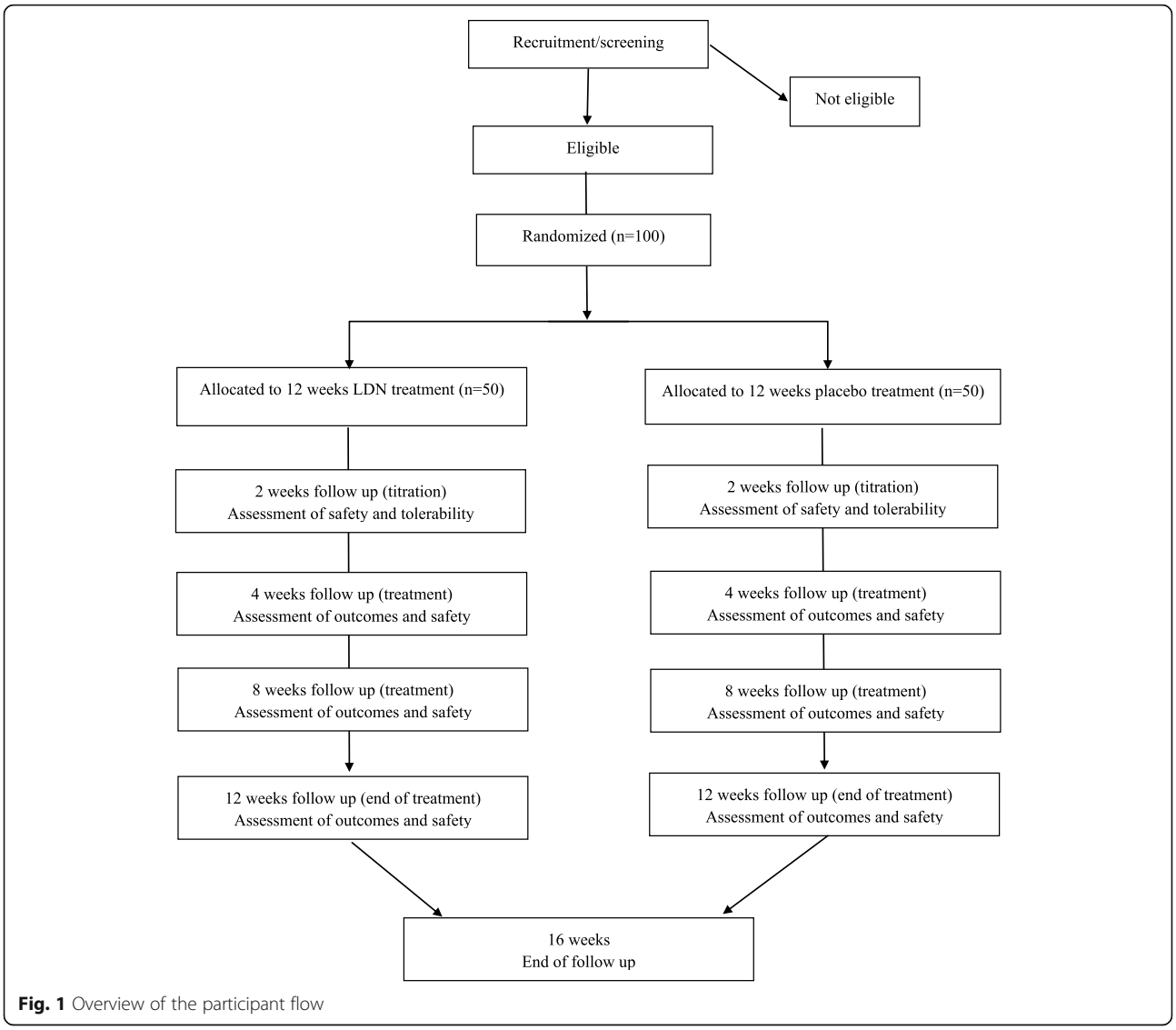


Fig. 1 Overview of the participant flow

and drop-out during the 12-week trial period, we decided to include 100 patients (with approximately 50 patients in each group), corresponding to a statistical power of more than 90% to detect a difference between groups in the ITT population.

If the intended sample size is not reached at 30 months after recruitment has started, the inclusion of patients will stop at 74 patients, which will ensure a power of 80%.

Recruitment {15}

Participants are recruited from a pain center at a public university hospital and through advertisement in relevant written and social/Web-based media. For ethical reasons, patients in active treatment at the pain center will not be recruited, but only patients who have completed treatment and signed up for participation in future medical trials or waiting list patients. To secure a broad representation of FM severity to the study population, recruitment through advertisement will be equally favored.

Assignment of interventions: allocation

Sequence generation {16a}

A computerized algorithm will be generated for randomization by preparing a list of 100 sequential numbers to active intervention or placebo intervention; randomization will be based on permuted blocks of 2–6 individuals. No stratifications are applied to the randomization, and both investigators and outcome assessors are blinded regarding the permuted blocking strategy.

Table 1 Schedule of enrolment, interventions, and assessments

	Study period						
	Enrolment	Allocation	Post-allocation				Follow-up
Week	*−4–0*	0	*2[a] (telephone)*	*4[a]*	*8[b]*	*12[b]*	*16[b]*
Enrolment							
Informed consent	X						
Medication history	X	X	X	X	X	X	X
Demographic data	X						
Eligibility screen	X						
Allocation		X					
Interventions							
Low-dose naltrexone							
Placebo							
Assessments							
Vital tests: blood pressure, weight, height		X				X	X
Safety tests: ALAT, creatinine, GFR, thrombocyte count, bilirubin. ECG	X					X	
hCRP		X					
Blood for biobank		X				X	
PROMs							
PHQ-9	X						
GAD-7	X						
FIQR		X		X	X	X	X
PGI-C				X	X	X	X
EQ-5D		X		X	X	X	X
EQ-VAS		X		X	X	X	X
Pain sensitivity							
Handheld algometry		X				X	
Computerized cuff algometry		X				X	
Muscle tests							
Isometric muscle exhaustion of deltoid		X				X	
30-s stand chair test		X				X	
Compliance assessment						X	
Adverse events		X	X	X	X	X	X

ALAT alanine aminotransferase, *GFR* glomerular filtration rate, *ECG* electrocardiogram, *PHQ-9* Patient Health Questionnaire – 9 items, *GAD-7* Generalized Anxiety Disorder – 7 items, *hCRP* high-sensitive C-reactive protein, *FIQR* Fibromyalgia Impact Questionnaire Revised, *PGI-C* Patient Global Impression of Change, *EQ-5D* EuroQol 5 dimensions, *EQ-VAS* EuroQol Visual Analogue Scale

[a]±2 days

[b]±7 days

Concealment mechanism {16b}

A data manager, with no clinical involvement in the trial, prepares the randomization sequence. The allocation is concealed in a password-protected computer file that is only accessible by the data manager. The randomization list is sent to the hospital pharmacy, who labels the medicine with blinding codes according to this list. The medicine is then shipped to the place of the trial. Unblinding will not take place before primary analysis of the data has taken place. In case unblinding of a single participant is necessary during the trial, individual allocations will also be held in sealed, opaque, consecutively numbered envelopes.

Implementation {16c}

The PI enrolls all participants. After signing the informed consent form, each participant is allocated a sequential number that randomizes them to one of the two groups.

Assignment of interventions: blinding

Who will be blinded {17a}

The study is triple-blind as participating patients, investigators, and outcome assessors (and statistical analysts) are blinded to the allocation. The active medicine and placebo tablets will look identical and will be blinded in similar cans and labeled with blinding codes.

Procedure for unblinding if needed {17b}

In the case of a suspected unexpected serious adverse reaction (SUSAR), the participant will be unblinded by the sponsor before reporting to the Danish Medicines Agency, but the PI will remain blinded. The PI will only be unblinded in the case of a medical emergency and only if the PI finds it necessary to ensure the safety of the subject. The PI can unblind a single subject by breaking the code-envelope for the subject's code number.

Data collection and management

Plans for assessment and collection of outcomes {18a}

After allocation has taken place, the participants will complete questionnaires at the beginning of every visit via an electronic survey and before talking to the investigators. The Fibromyalgia Impact Questionnaire Revised [22] is a disease-specific instrument, while the EQ-5D-5L (which includes the EQ-VAS) [25] is a generic instrument. All are validated for use in clinical trials.

The level of tenderness is assessed at baseline and after 12 weeks of treatment using a handheld pressure algometer (Somedic Algometer, Hørby, Sweden). Assessment sites are the right quadriceps muscle 15 cm from the apex patella and the left trapezius muscle 10 cm from acromion (between acromion and C6/7). Each site is assessed three times. To avoid bias due to interrater variability, the same investigator will carry out all the procedures.

The exploratory outcome measures are assessed by an independent assessor at baseline and after 12 weeks of treatment. Standard operating procedures will be available, and the assessor will be trained in the procedures before and during the trial. The procedures are:

- Computer-controlled cuff algometry on lower legs in all participants to assess pressure pain threshold, pressure pain tolerance, temporal summation of pain, and conditioned pain modulation. Standardized assessment of experimental pressure pain sensitivity has shown good reliability and provides insights into the pathophysiological mechanisms involved in the pain condition.
- Muscular exhaustion: the participant completes an isometric muscle exhaustion task by maintaining 90° shoulder abduction (dominant arm) for as long as possible with the elbow extended and the hand pronated (hand facing downwards). Task failure (test position can no longer be maintained) defines the test duration. Surface electromyography (EMG) will be recorded from the anterior, middle, and posterior deltoid muscle at 3000 Hz during the entire test. The test has been shown to be feasible in women with fibromyalgia [26].
- Physical fitness is measured by the 30-s chair stand test, which has been shown to be reliable and feasible in women with fibromyalgia [27].

Plans to promote participant retention and complete follow-up {18b}

The participants will receive a daily short text message (SMS) reminding them to take their trial medication. Participants who discontinue the treatment during the trial will be encouraged to complete all visits as scheduled.

Data management {19}

The participants enter questionnaire data directly via a survey into the electronic Case Report File (eCRF) using REDCap electronic data capture tools. The EMG files are saved in a secured and logged Sharepoint. Results from the protocol-specific procedures will be collected in paper format and then entered into the eCRF. The assessors enter all other data directly into the eCRF during the visits. Data quality in the eCRF will be promoted using range checks for data values. Data will later be transferred to a statistical program for analyses. The data will be anonymized 5 years after the termination of the study.

Confidentiality {27}

All data about potential and enrolled participants will be collected in a secure and logged database, in a secure and logged Sharepoint, or behind a double lock for data in paper format. Only anonymized data will be shared.

Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in this trial/future use {33}

Blood for a research biobank will be collected before baseline and after 12 weeks of treatment. The purpose of the biobank is to be able to measure a possible change in pro- and anti-inflammatory cytokines in participants receiving active treatment compared to placebo. For this purpose, 2 × 0.5 ml serum and 2 × 0.5 ml plasma are collected before baseline and after 12 weeks of treatment. Any excess blood will be stored for 10 years. After 10

years, the blood will be destroyed. Informed consent to perform analyses for other research purposes is collected from all participants.

Statistical methods

Statistical methods for primary and secondary outcomes {20a}

The main analyses will be based on the intention-to-treat (ITT) population. This ITT principle asserts the effect of a treatment policy (that is, the planned treatment regimen) rather than the actual treatment given (i.e., it is independent of treatment adherence). Accordingly, participants allocated to a treatment group at baseline (X_{LDN} or $X_{Placebo}$) will be followed up, assessed, and analyzed as members of that group, irrespective of their adherence to the planned course of treatment (i.e., independent of withdrawals and cross-over phenomena). By using mixed effects models (explained below), missing data after baseline will be handled indirectly; mixed effects models are valid assuming data are "Missing at Random" (MAR) [28].

All *P* values and 95% confidence intervals (*95% CI*) will be two-sided. We will not apply explicit adjustments for multiplicity; rather, we will analyze and interpret the 21 secondary outcomes in a prioritized order (e.g., "gatekeeping procedure" and/or the Hochberg sequential procedure). The analyses of the key secondary outcomes will be performed in sequence until one of the analyses fails to show the statistically significant difference, or until all analyses have been completed at a statistical significance level of 0.05 (i.e., *95% CI* does not overlap "the null").

Unlike the Bonferroni correction/interpretation (directly adjusting the statistical significance threshold by the number of tests planned [say, k] $\rightarrow P^* = 0.05/k$), we will apply the Hochberg sequential procedure, where all the tests are performed and the resultant *P* values are ordered from largest to smallest on a list [29]. With our statistical significance level fixed at 5% and the largest observed if the *P* value is less than .05, then all the tests will be considered significant. Otherwise, if the next largest *P* value is less than 0.05/2 (.025), then all the tests except the one with the largest *P* value are considered significant. This process will be continued until all the comparisons made have been interpreted. This approach uses progressively more stringent statistical thresholds with the most stringent one being the Bonferroni threshold. This approach will achieve a greater power to detect true effect than the Bonferroni procedure [30].

Our primary (main) analyses will be based on the estimation 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) are used based on a linear mixed model where the treatment group is used as a fixed effect and participant ID as a random-effect parameter. All between-group differences will be adjusted for baseline level in order to reduce the random variation. The primary statistical model will consist of fixed effects and random effects. Fixed effects define the expected values of the observations, and random effects define the variance and covariances of the observations. In this study, participants will be randomly assigned to two treatment groups (X_{LDN} vs $X_{Placebo}$), and observations are made at four time points for the primary outcome measure (baseline and 4, 8, and 12 weeks from baseline). Basically, there are two fixed-effect factors: group and time. Random effects result from variation between and within participants. We anticipate that measures on the same patient at different times are correlated, with measures taken closely together in time being more highly correlated than measures taken more apart in time. Observations on different participants will be assumed to be independent.

Secondarily, an analysis of the number of responders (dichotomous outcomes) in the two groups will be carried out using logistic regression analyses. A responder is defined as a participant who reports a more than 15%, 30%, or 50% decrease in pain after 12 weeks of treatment with LDN. For these dichotomous outcomes, logistic regression will be used to calculate the odds ratio (*OR*) with *95% CI* comparing the two groups. For subsequent ease of interpretation, the *OR* values will be converted into (relative) risk ratios and (absolute) risk differences. The pre-specified efficacy analyses will be based on the data for the full analysis set, the ITT population, which includes all participants assessed and randomized at baseline.

Interim analyses {21b}

Not applicable as no interim analysis is made.

Methods for additional analyses (e.g., subgroup analyses) {20b}

Not applicable as no subgroup analyses are made.

Methods in analysis to handle protocol non-adherence and any statistical methods to handle missing data {20c}

Repeated measurements using mixed models will be based on the ITT population, including all randomized participants with available data at baseline. Missing data will be handled indirectly and statistically modeled using repeated-measures linear mixed models (see below). These models will be valid if data are missing at random (MAR): "Any systematic difference between the missing values and the observed values can be explained by differences in observed data" [28]. Contrasts between

groups will be estimated based on repeated-measures analysis of covariance applied in mixed linear models (at 12 weeks from baseline). Thus, in the case of missing data during the 12-week trial, repeated-measures linear mixed models will adjust for that indirectly.

To confirm the robustness of the findings for the primary and key secondary outcomes, sensitivity analyses will be performed on the main analyses including the:

(i) "Complete Case" population, i.e., outcome data recorded both at baseline and after 12 weeks; a dataset potentially valid if data are missing completely at random (MCAR)
(ii) Non-responder imputation: use of single imputation where the baseline observation is carried forward; potentially valuable if data are not missing at random (NMAR)
(iii) "Per Protocol" population: defined as participants with at least 80% adherence to treatment

Robustness is a concept that refers to the sensitivity of the overall conclusions to various limitations of the data, assumptions, and analytic approaches to data analysis. Robustness implies that the treatment effect and primary conclusions of the FINAL trial are not substantially affected when analyses are carried out based on alternative assumptions or analytic approaches.

Plans to give access to the full protocol, participant-level data, and statistical code {31c}

The full protocol and the statistical analysis plan (SAP) will be accessible at www.clinicaltrials.gov, identifier: NCT04270877.

Oversight and monitoring

Composition of the coordinating center and trial steering committee {5d}

Not applicable as it is a single-center study.

Composition of the data monitoring committee, its role, and reporting structure {21a}

The Good Clinical Practice (GCP) unit at Odense University Hospital monitors the trial.

Adverse event reporting and harms {22}

Data on adverse events (AEs) and adverse reactions (ARs) are collected at all visits. The participants will complete a questionnaire about the presence of known side effects and will be interviewed by the PI about any adverse events that occur during the trial. For the graduation of the severity of harms, the Common Terminology Criteria for Adverse Events (CTCAE) version 5.0 will be used. The PI assesses whether an AE is related to the trial medication using the Summary of Product Characteristics (SmPC) for Naltrexone 50 mg as a reference document. All AEs and ARs are described in detail and registered in the eCRF.

ALAT, bilirubin, creatinine, GFR, thrombocyte count, and electrocardiogram are assessed before and after the intervention. Urinary human chorionic gonadotropin is measured at baseline (week 0) and after 4, 8, and 12 weeks of treatment in all women of fertile age.

A serious adverse event (SAE) is any untoward medical occurrence or effect that at any dose results in death, is life-threatening, requires hospitalization or prolongation of existing hospitalization, results in persistent or significant disability or incapacity, or is a congenital anomaly or birth defect. All SAEs are reported by the PI to the sponsor within 24 h. Causality of an SAE will be determined according to the detailed guidance on the collection, verification, and presentation of adverse event/reaction reports arising from clinical trials on medicinal products for human use (CT-3) guidelines. If a serious adverse reaction (SAR) is assessed as unexpected according to the SmPC, the sponsor must unblind the subject before reporting it to the Danish Medicines Agency. The PI will remain unblinded. Under section 89 [2](i) of the Danish Medicines Act, the sponsor must immediately inform the Danish Medicines Agency if any SUSARs occur during the trial.

Frequency and plans for auditing trial conduct {23}

Not applicable as no auditing.

Plans for communicating important protocol amendments to relevant parties (e.g., trial participants, ethical committees) {25}

Any important protocol amendments will be reported to the Danish Medicines Agency, the local ethical committee, the monitor of the trial, participating investigators, trial participants, relevant trial registries, and the journal that has published the protocol.

Dissemination plans {31a}

Information about the trial is published at ClinicalTrials.gov and the European Union Drug Regulating Authorities Clinical Trials Database (EUDRACT) before enrolment of the first patient. The protocol and study results will be published in international peer-reviewed journals. Both positive, negative, and inconclusive results will be published. After publication, the results from the trial will be disseminated to the trial participants via email and to the public via written and Internet media.

Discussion

The traditional pharmacological treatment of chronic non-malignant pain (CNMP), which includes FM, aims at reducing facilitatory neurotransmitters (e.g., gabapentinoids) or increasing inhibitory neurotransmitters (e.g., serotonin-norepinephrine reuptake inhibitors) [31]. These treatments do not always result in satisfactory pain relief, however, and their use is often limited by side effects. Furthermore, traditional therapies do not necessarily offer relief from other key symptoms associated with CNMP/FM. The results of our previous dose-response study indicated that LDN has a positive influence on sleep disturbance, energy, and touch tenderness in women with FM [18]. This is in concordance with previous trials on efficacy [14, 15]. Thus, treatment with LDN might offer several advantages to existing treatments such as new targets of action, fewer side effects, and a relatively low cost.

Currently, LDN is widely used as an off-label treatment for CNMP including FM, but the evidence is based on case reports and a few small clinical trials. This will be the first high-quality trial of LDN with a sufficient sample size to investigate a clinically relevant change in pain in women with FM. In addition, the current randomized, placebo-controlled trial aims to provide high-quality evidence by reducing the risk of bias through blinding of participating patients, investigators, outcome assessors, and statistical analysts. Finally, the transparency of the applied methods and definitions of outcome measures will be ensured through public access to the current protocol paper and a priori registration at ClinicalTrials.gov.

This trial contains both pragmatic and exploratory elements. The study will be the first to explore the efficacy of 6 mg LDN in women with fibromyalgia. However, for pragmatic reasons, a titration phase allows the testing of lower doses in case of problems with tolerability, aiming to assess the effectiveness of LDN on pain and other FM symptoms. Another pragmatic attitude is to secure a broad spectrum of FM severity recruiting participants through advertising and allowing for continued use of different kinds of usual care. The inclusion of exploratory outcomes aims to examine the mechanisms of action of LDN. If an effect of LDN on pain sensitivity, muscle fatigue, or biomarkers for CNS inflammation can be demonstrated in women with FM, this will not only expand our knowledge about mechanisms of action of LDN but might also contribute to a better understanding of underlying pathology in FM.

FM represents a well-defined subgroup of CNMP that is suitable for clinical trials. The disorder is characterized by chronic widespread pain and widespread hyperalgesia to mechanical stimulation [19] and is a classic example of a nociplastic pain disorder hypothesized to be caused primarily by disturbances in central pain regulatory mechanisms [32, 33]. Findings from trials in FM patients might therefore be extrapolated to other primary pain conditions with nociplastic pain features. As FM is diagnosed more frequently in women [34], recruitment of men with FM can be difficult. We have therefore chosen to include only women in order to ease recruitment and strengthen the internal validity of the results. This will have an impact on generalizability and external validity, and the final results must be reproduced later in a population including men.

Trial status

Protocol version 5.1. Date: 27.07.2021 (dd.mm.yyyy)
Approval from authorities: 30.10.2019
Expected start of inclusion: 01.11.2020
Expected end of inclusion: 01.01.2023
Expected end of follow-up:01.06.2023

Abbreviations

LDN: Low-dose naltrexone; FM: Fibromyalgia; NLX: Naltrexone; CNS: Central nervous system; TLR4: Toll-like receptor 4; RCT: Randomized controlled trial; NRS: Numeric rating scale; ALAT: Alanine aminotransferase; GFR: Glomerular filtration rate; PHQ-9: Patient Health Questionnaire – 9 items; PI: Primary investigator; SMS: Short text message; OMERACT: Outcome Measures in Rheumatology Clinical Trials; FIQR: Fibromyalgia Impact Questionnaire Revised; WPI: Widespread Pain Index; EQ-VAS: EQ-5D visual analog scale; CPA: Computerized pressure cuff algometry; CPM: Conditioned pain modulation; TSP: Temporal summation of pain; IMMPACT: The Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials; MCID: Minimal clinical important difference; ITT: Intention-to-treat; SUSAR: Suspected unexpected serious adverse reaction; EMG: Surface electromyography; eCRF: Electronic Case Report File; *CI*: Confidence interval; *OR*: Odds ratio; MAR: Missing at random; MCAR: Missing completely at random; NMAR: Not missing at random; SAP: Statistical analysis plan; GCP: Good Clinical Practice; AE: Adverse event; AR: Adverse reaction; SmPC: Summary of Product Characteristics; SAE: Serious adverse event; CT-3: The detailed guidance on the collection, verification, and presentation of adverse event/reaction reports arising from clinical trials on medicinal products for human use; SAR: Serious adverse reaction; EUDRACT: European Union Drug Regulating Authorities Clinical Trials Database; CNMP: Chronic non-malignant pain

Acknowledgements

Study data will be collected and managed using REDCap electronic data capture tools hosted at Open Patient data Explorative Network (OPEN), Odense University Hospital, Region of Southern Denmark.
The authors thank Claire Gudex, Department of Clinical Research, University of Southern Denmark, for editing the manuscript.

Authors' contributions {31b}

KDB is the primary investigator; she conceived the study and led the proposal and protocol development. KA contributed to the study design and protocol development and was the lead specialist in fibromyalgia. HBV contributed to the study design and protocol development and was the lead specialist in the assessment of pain sensitivity. MRBE contributed to the study design and protocol development and was the lead specialist in biobanks and the measurement of cytokines. AHL contributed to the study design and protocol development and was the lead specialist in the measurement of muscle exhaustion and physical fitness. RC contributed to the study design and protocol development and was the lead trial statistician. PT is the sponsor of the trial and was the lead trial methodologist. All authors read and approved the final manuscript.

Funding {4}
The Danish Rheumatism Association has granted 175,000 DKK for operating expenses. Odense University Hospital's PhD fund for operation expenses has granted 100,000 DKK towards operating expenses and the study nurse's salary. Aase and Ejnar Danielsens fund has granted 100.000 DKK towards operating expenses and the study nurse's salary. The Parker Institute at Bispebjerg and Frederiksberg Hospital is supported by a core grant from the Oak Foundation (OCAY-18-774-OFIL). The work of the PI is covered by a PhD salary.
None of the funders plays any role in the design of the study or in the collection, analysis, and interpretation of data or in writing the manuscript.

Availability of data and materials {29}
The trial is investigator initiated, and only the investigators have access to the final trial dataset and the material from the biobank.

Declarations

Ethics approval and consent to participate {24}
Approval has been obtained from the Ethical Committee of Southern Denmark (S-20190133). Written informed consent to participate will be obtained from all participants.

Consent for publication {32}
Not applicable.

Competing interests {28}
The authors declare that they have no competing interests.

Author details
[1]Pain Research Group, Pain Center, Odense University Hospital, Heden 7-9, Indgang 20, DK – 5000 Odense C, Denmark. [2]Department of Clinical Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark. [3]The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark. [4]Department of Rheumatology, Bispebjerg and Frederiksberg Hospital, Frederiksberg, Denmark. [5]Department of Orthopedics and Traumatology, Odense University Hospital, Odense, Denmark. [6]Department of Anesthesiology and Intensive Care, Odense University Hospital, Odense, Denmark.

Received: 6 April 2021 Accepted: 29 October 2021
Published online: 15 November 2021

References

1. Patten DK, Schultz BG, Berlau DJ. The safety and efficacy of low-dose naltrexone in the management of chronic pain and inflammation in multiple sclerosis, fibromyalgia, Crohn's disease, and other chronic pain disorders. Pharmacotherapy. 2018;38(3):382–9. https://doi.org/10.1002/phar.2086.
2. Yoon G, Kim SW, Thuras P, Westermeyer J. Safety, tolerability, and feasibility of high-dose naltrexone in alcohol dependence: an open-label study. Hum Psychopharmacol. 2011;26(2):125–32. https://doi.org/10.1002/hup.1183.
3. Verebey K. The clinical pharmacology of naltrexone: pharmacology and pharmacodynamics. NIDA Res Monogr. 1981;28:147–58.
4. Valenta JP, Job MO, Mangieri RA, Schier CJ, Howard EC, Gonzales RA. mu-opioid receptors in the stimulation of mesolimbic dopamine activity by ethanol and morphine in Long-Evans rats: a delayed effect of ethanol. Psychopharmacology (Berl). 2013;228(3):389–400.
5. Crabtree BL. Review of naltrexone, a long-acting opiate antagonist. Clin Pharm. 1984;3(3):273–80.
6. Gillman MA, Lichtigfeld FJ. A pharmacological overview of opioid mechanisms mediating analgesia and hyperalgesia. Neurol Res. 1985;7(3):106–19. https://doi.org/10.1080/01616412.1985.11739709.
7. Tempel A, Kessler JA, Zukin RS. Chronic naltrexone treatment increases expression of preproenkephalin and preprotachykinin mRNA in discrete brain regions. J Neurosci. 1990;10(3):741–7. https://doi.org/10.1523/JNEUROSCI.10-03-00741.1990.
8. Younger J, Parkitny L, McLain D. The use of low-dose naltrexone (LDN) as a novel anti-inflammatory treatment for chronic pain. Clin Rheumatol. 2014;33(4):451–9. https://doi.org/10.1007/s10067-014-2517-2.
9. Loggia ML, Chonde DB, Akeju O, Arabasz G, Catana C, Edwards RR, et al. Evidence for brain glial activation in chronic pain patients. Brain. 2015;138(Pt 3):604–15. https://doi.org/10.1093/brain/awu377.
10. Bruno K, Woller SA, Miller YI, Yaksh TL, Wallace M, Beaton G, et al. Targeting toll-like receptor-4 (TLR4)-an emerging therapeutic target for persistent pain states. Pain. 2018;159(10):1908–15. https://doi.org/10.1097/j.pain.0000000000001306.
11. Ghai B, Bansal D, Hota D, Shah CS. Off-label, low-dose naltrexone for refractory chronic low back pain. Pain Med. 2014;15(5):883–4. https://doi.org/10.1111/pme.12345.
12. Hota D, Srinivasan A, Dutta P, Bhansali A, Chakrabarti A. Off-label, low-dose naltrexone for refractory painful diabetic neuropathy. Pain Med. 2016;17(4):790–1. https://doi.org/10.1093/pm/pnv009.
13. Sturn KM, Collin M. Low-dose naltrexone: a new therapy option for complex regional pain syndrome type I patients. Int J Pharm Compd. 2016;20(3):197–201.
14. Younger J, Mackey S. Fibromyalgia symptoms are reduced by low-dose naltrexone: a pilot study. Pain Med. 2009;10(4):663–72. https://doi.org/10.1111/j.1526-4637.2009.00613.x.
15. 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(2):529–38. https://doi.org/10.1002/art.37734.
16. Parkitny L, Younger J. Reduced pro-inflammatory cytokines after eight weeks of low-dose naltrexone for fibromyalgia. Biomedicines. 2017;5(2):16.
17. Bolton MJ, Chapman BP, Van Marwijk H. Low-dose naltrexone as a treatment for chronic fatigue syndrome. BMJ Case Rep. 2020;13(1):e232502.
18. 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(10):2253–61. https://doi.org/10.1093/pm/pnaa001.
19. Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, et al. The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multicenter Criteria Committee. Arthritis Rheum. 1990;33(2):160–72. https://doi.org/10.1002/art.1780330203.
20. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606–13. https://doi.org/10.1046/j.1525-1497.2001.016009606.x.
21. Mease PJ, Clauw DJ, Christensen R, Crofford LJ, Gendreau RM, Martin SA, et al. Toward development of a fibromyalgia responder index and disease activity score: OMERACT module update. J Rheumatol. 2011;38(7):1487–95. https://doi.org/10.3899/jrheum.110277.
22. 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(4):R120. https://doi.org/10.1186/ar2783.
23. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Hauser W, Katz RL, et al. 2016 Revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum. 2016;46(3):319–29. https://doi.org/10.1016/j.semarthrit.2016.08.012.
24. Dworkin RH, Turk DC, McDermott MP, Peirce-Sandner S, Burke LB, Cowan P, et al. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. Pain. 2009;146(3):238–44. https://doi.org/10.1016/j.pain.2009.08.019.
25. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: past, present and future. Applied Health Economics and Health Policy. 2017;15(2):127–37. https://doi.org/10.1007/s40258-017-0310-5.
26. Bandak E, Amris K, Bliddal H, Danneskiold-Samsoe B, Henriksen M. Muscle fatigue in fibromyalgia is in the brain, not in the muscles: a case-control study of perceived versus objective muscle fatigue. Ann Rheum Dis. 2013;72(6):963–6. https://doi.org/10.1136/annrheumdis-2012-202340.
27. Martin-Martinez JP, Collado-Mateo D, Dominguez-Munoz FJ, Villafaina S, Gusi N, Perez-Gomez J. Reliability of the 30 s chair stand test in women with fibromyalgia. Int J Environ Res Public Health. 2019;16(13):2344.
28. Detry MA, Ma Y. Analyzing repeated measurements using mixed models. Jama-J Am Med Assoc. 2016;315(4):407–8. https://doi.org/10.1001/jama.2015.19394.
29. Cao J, Zhang S. Multiple comparison procedures. JAMA. 2014;312(5):543–4. https://doi.org/10.1001/jama.2014.9440.
30. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. Behav Brain Res. 2001;125(1-2):279–84. https://doi.org/10.1016/S0166-4328(01)00297-2.
31. Clauw DJ. Fibromyalgia: a clinical review. JAMA. 2014;311(15):1547–55. https://doi.org/10.1001/jama.2014.3266.

32. Yunus MB. Central sensitivity syndromes: a new paradigm and group nosology for fibromyalgia and overlapping conditions, and the related issue of disease versus illness. Semin Arthritis Rheum. 2008;37(6):339–52. https://doi.org/10.1016/j.semarthrit.2007.09.003.
33. Gracely RH, Ambrose KR. Neuroimaging of fibromyalgia. Best Pract Res Clin Rheumatol. 2011;25(2):271–84. https://doi.org/10.1016/j.berh.2011.02.003.
34. Wolfe F, Walitt B, Perrot S, Rasker JJ, Hauser W. Fibromyalgia diagnosis and biased assessment: sex, prevalence and bias. PLoS One. 2018;13(9):e0203755. https://doi.org/10.1371/journal.pone.0203755.

Publisher's Note

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.

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

Lancet Rheumatol 2023

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

See Online/Comment https://doi.org/10.1016/S2665-9913(23)00297-7

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**

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.

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.

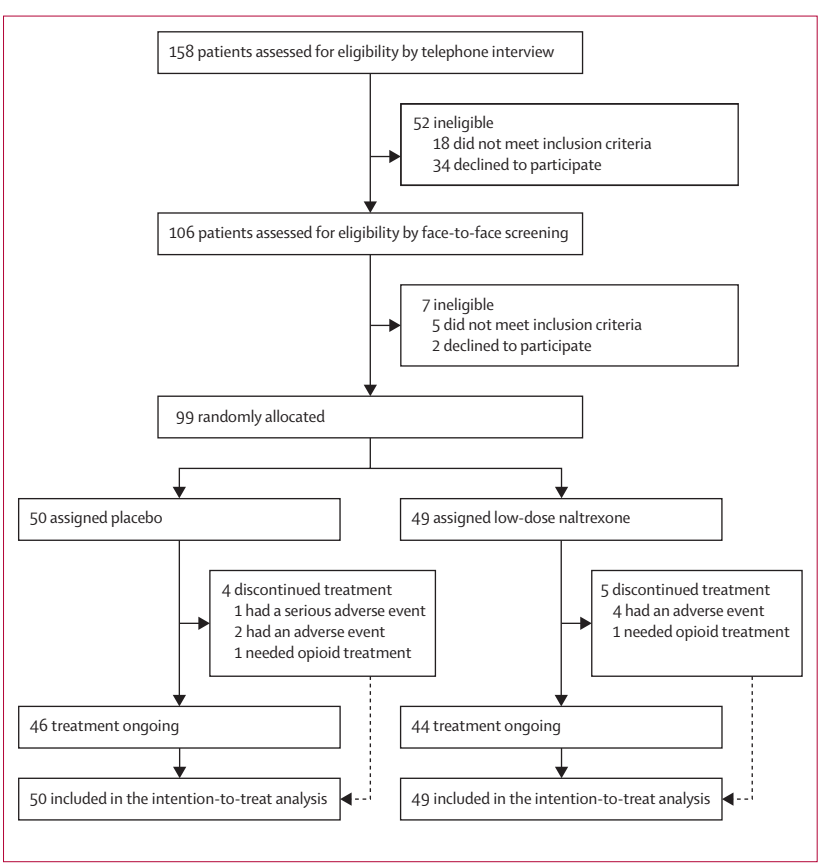


Figure 1: **Trial profile**

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 treatment. No participants were lost to follow-up, and the

	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**

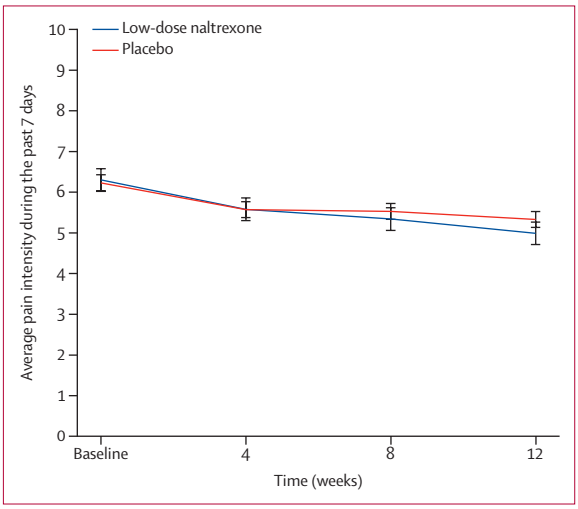


Figure 2: **Pain trajectory**
The trajectory for the average pain during the past 7 days (using the pain question from the Fibromyalgia Impact Questionnaire, revised) over time from baseline to the primary endpoint after 12 weeks.

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.