UiT The Arctic University of Norway

Faculty of health science, Department of community medicine

Do patients assigned to multidisciplinary examination differ from patients assigned to monodisciplinary examination at the University Hospital of Northern Norway?

The Norwegian neck and back registry 2018

Kjetil Magne Samuelsen

Master's thesis in Public Health HEL-3950 December 2020

Supervisor: Kristin Benjaminsen Borch Co-supervisor: Maja Wilhelmsen

Foreword

I would like to thank my supervisor, Kristin Benjaminsen Borch, for all help through this long process. I have truly learned a lot during our many conversations about this thesis and research in general. I also appreciate your kind and friendly demeanor, which helped keeping my spirits up when I wasn't always feeling too good about my work. I was very lucky when you offered to be my supervisor, Kristin.

I would also thank my employer, the University Hospital of Northern Norway, for having faith in me and granting me this opportunity to expand my knowledge in the field of public health. My colleagues at the HelseIArbeid-clinic have gotten accustomed to me coming and going over the last 2,5 years, but I have always felt great support along the way. I look forward to returning full-time and hopefully I can help our clinic continue to evolve through my education.

I would also thank my co-supervisor, Maja Wilhelmsen, especially for helping me in the initial phase of this thesis navigating the application process of obtaining data from the NNRR.

Most of all I would like to thank my beautiful partner, Tone Mathilde, for all your dearly needed support the last 2,5 years. These last couple of years have been a whirlwind. We have started a family in the same period, and without your dedication this process would have been a lot harder. I look forward to returning to a normal everyday now with you and Aasmund. You two are truly the best.

Table of Contents

Abstract

Background: Back pain is a very common disability worldwide and in Norway. The term "back pain" refers to all pain from the back; and includes neck pain and low back pain (LBP), which are the two most common parts of the back to experience pain from. 60-80% of the population in Norway will experience LBP during their life. 30-50% of all people report neck pain during a year. A large proportion back pain has no pathoanatomical diagnosis. Back pain is considered a multifactorial condition where biopsychosocial factors influence the patient's pain. The term "yellow flags" are commonly used in clinical guidelines and research when referring to psychosocial risk factors for developing chronic pain. The biopsychosocial understanding of back pain has led to the development of multidisciplinary interventions. The Norwegian clinical guideline of LBP recommends referral of patients to a multidisciplinary rehabilitation program like a "specialized clinic in physical medicine and rehabilitation" if pain persist after 6-8 weeks. The guidelines also define a multidisciplinary team doing examinations at a specialized clinic as a team of at least two health professionals. No studies have previously investigated the patients assigned to multidisciplinary versus monodisciplinary examinations at a specialized clinic in Norway.

Aims: Investigate if the patients assigned to multidisciplinary examination differ from patients assigned to monodisciplinary examination regarding patient characteristics, yellow flags, pain indicators, function level and treatment recommendations.

Material and methods: Data was extracted from the Norwegian Neck and Back Registry (NNRR). All patients who were examined at the specialized clinic for physical medicine and rehabilitation at the University Hospital of Northern Norway (UNN) in 2018 and completed the first patient questionnaire of the NNRR were included in the original dataset. After exclusions, the final study sample consisted of 655 patients. The differences between the patients assigned to either mono- or multidisciplinary examinations were tested using the Chi-Square test and the Independent Samples T-Test. Binary logistic regression was used to calculate multivariable adjusted odds ratios (OR) with 95% confidence intervals (95% CI) of associations between exposure variables and being assigned to mono- or multidisciplinary examinations.

Results: Patients assigned to multidisciplinary examinations were on average 5.6 years younger. Among the yellow flags describing mental health problems there were increased

odds for the multidisciplinary group to report both depression (OR=2.04, 95% CI=1.09-3.80), anxiety (2.06, 95% CI=1.03-4.12) and a HSCL-10 score >1.85 indicating mental health problems (OR=1.64, 95% CI=0.96-2.77) compared to patients assigned to monodisciplinary examinations. There were also increased odds (OR=2.53, 95% CI=1.25-5.10) that the patient believed his/her pain was caused by mental problems. Among yellow flags describing comorbidities there were increased odds of stomach discomfort (OR=2.16, 95% CI=1.10-4.24), upper back pain (OR=1.74 95% CI=1.04-2.90) and shoulder pain (OR=1.85 95% CI=1.12-3.06) in the multidisciplinary group compared to the monodisciplinary group. There were increased odds (OR=2.36, 95% CI=1.20-4.64) of the patients in the multidisciplinary group using prescription pain medication more than once every week compared to the patients in the monodisciplinary group. Patients in the multidisciplinary group had lower odds (OR=0.43, 95% CI=0.24-0.76) of being recommended no treatment compared to the patients in the monodisciplinary group, while at the same time having increased odds (OR=3.1, 95% CI=1.85-5.20) of being recommended treatment for follow-up by a physician in primary care.

Conclusion: This study shows that patients assigned for multidisciplinary examinations at UNN are younger than patients assigned for monodisciplinary examinations. There are also indications that patients assigned to multidisciplinary examinations suffers from more mental health problems, as well as being more likely to use pain medication requiring a prescription. The patients assigned to multidisciplinary examinations were also receiving overall more treatment recommendations from the health professionals and were especially more likely to be recommended to follow-up by a physician in primary care. There were no differences between the groups concerning pain, function and work-related variables.

Abbreviations

CI	Confidence Interval
DALY	Disability Adjusted Life Year
FAB-Q	Fear-Avoidance Beliefs Questionnaire
GP	General Practitioner
HSCL-10	Hopkins Symptoms Check List 10
ICD-10	International Classification of Diseases 10th Revision
LBP	Low Back Pain
NDI	Neck Disability Index
NNRR	Norsk Nakke- og Ryggregister – Norwegian Neck and Back Registry
NPGD	National Priority Guidance Document
ODI	Oswestry Disability Index
OR	Odds Ratio
PROMs	Patient Reported Outcome Measures
SHC	Subjective Health Complaints
UHI	Ursin Health Inventory
UNN	University Hospital of Northern Norway
VAS	Visual Analogue Scale

1 Introduction

1.1 Back pain

Low back pain (LBP) is a very common disability both in the global (1) and Norwegian population (2). Globally it is estimated that 23.2% of the population experience LBP during any one-month period. It is estimated that 60-80% of the population in Norway will experience LBP during their life. 30-50% of all people also report neck pain during a year (3, 4). Throughout this text when addressing pain from both the neck and low back, the term "back pain" is used. Furthermore, LBP and neck pain are the largest reasons for loss of Disability Adjusted Living Years (DALY) (5, 6). This means that back pain is a very common cause of personal suffering, as well as being a very costly patient group due to extensive sick leave and disability welfare in Norway (2, 7).

LBP is often referred to as "non-specific LBP" as it is almost always impossible to find a specific cause of the back pain (8). It is estimated that 85% of patients with LBP have nonspecific LBP with no pathoanatomical diagnosis (9, 10). As such LBP is considered as a multifactorial condition where biophysical factors, psychological factors, social and societal factors (biopsychosocial), and central pain processing and modulation influence the pain that the patient experience (8). Therefore, LBP is to be understood as a symptom and not a specific disease.

Neck pain is also most commonly non-specific as only a very small fraction of new episodes are caused by nerve irritation or major structural pathology (11). Neck pain is also associated with biopsychosocial factors and co-morbidities like headache and other musculoskeletal disorders (4).

In the literature additional risk factors for developing persistent or chronic pain are referred to as "yellow flags" (10). Effective treatment and interventions aimed at LBP is scarce (12), though multidisciplinary interventions have been suggested to be more effective for pain patients in general (13) and for LBP patients (12, 14).

1.1.1 Multidisciplinary interventions

Both neck and LBP are associated with psychological and social factors and this has led to the development of interventions aimed at tackling multiple factors and involving multiple types of health professionals (15).

Several countries have clinical guidelines for treatment of LBP patients, including Norway (10). In an overview of 15 clinical guidelines for management of non-specific LBP 11 of 15 guidelines recommend some type of multidisciplinary rehabilitation (16). In the Norwegian clinical guideline of LBP it is recommended to refer patients to a multidisciplinary rehabilitation program like a "specialized clinic in physical medicine and rehabilitation" if pain persist after 6-8 weeks (10). There is no existing clinical guideline for treatment of neck pain in Norway, but a multidisciplinary approach has been suggested as important for many of the patients with neck pain (4).

There is though no clear description of what a multidisciplinary intervention should consist of other than being performed by more than one health profession. As a multidisciplinary intervention would have several different components, it is not clear what components are responsible for better patient outcomes due to multidisciplinary interventions. Little is known about the duration, setting or type of multidisciplinary intervention that could be more effective (14). The European Guidelines for Prevention of Back Pain also notes the problem of defining "multidisciplinary programs" (17).

In the clinical guidelines for treatment of LBP in Norway a multidisciplinary team doing examinations at a specialized clinic in physical medicine and rehabilitation are defined as a team of at least two health professionals (10). The multidisciplinary examinations at the specialized clinic of physical medicine and rehabilitation at the University Hospital in Northern Norway (UNN) follow this definition.

1.1.2 Yellow flags – risk factors for developing chronic back pain

In the transition from the earlier biomedical view of back pain as a disease to the now accepted view of back pain as a multifactorial and biopsychosocial symptom (8), a lot of effort has gone into identifying risk factors or prognostic predictors for developing chronic back pain. The definition of chronic back pain is pain persisting over twelve weeks or 3 months (10, 18). Chronic pain is also commonly referred to as persistent pain or long-lasting pain. Psychosocial factors have been showed to be risk factors in developing chronic pain

(19-23). The term “yellow flags” are commonly used in both clinical guidelines (10, 17) and research when referring to psychosocial risk factors. The term was originally created to capture the psychological risk factors and social and environmental risk factors for developing persistent disability due to musculoskeletal symptoms (23).

In the Norwegian clinical guideline for LBP yellow flags are listed as risk factors for developing chronic back pain, mainly being psychosocial factors (10):

- Work related issues and sick leave.
- Mental health problems like anxiety or depression.
- Comorbidities like generalized pain, headache, tiredness, dizziness and stomach issues.
- Past debilitating back pain potentially with affection of nerves.
- Pessimistic or negative attitude towards the pain. Fear-avoidance beliefs towards physical activity or work. Low expectation of recovery or return to work.

In this study it was chosen to use these points from the Norwegian guideline as an overview when assessing differences among yellow flags between patients assigned to either mono- or multidisciplinary examinations.

1.2 Back pain intervention at hospitals

To understand the course of the patient with back pain after referral from the primary care to the hospitals in Norway the following gives a short description. The larger hospitals in Norway have multidisciplinary specialized clinics in physical medicine and rehabilitation who examine and treat back pain patients who are referred from primary care. In most cases these patients have non-specific neck pain or non-specific LBP. Other back pain patients with suspected specific pathology are in general referred to other clinics at the hospital like a neurosurgical, rheumatological or orthopedic clinic. The specialized clinics in physical medicine and rehabilitation are required to follow national clinical guidelines, but they still organize themselves differently with regards to available staff and services provided. Commonly the health professionals working at specialized clinics in physical medicine and rehabilitation are senior physicians, physicians in specialization and physiotherapists. Other health professionals available could be psychologists and occupational therapists.

The patients are in general referred to the clinics by their general practitioners (GP)(Figure 1), but can also be referred from other hospital clinics or chiropractors and manual therapists in primary care. The referral is initially reviewed by one of the senior physicians working at the clinic. The senior physician then assigns the patients to an examination or rejects the case based on information given in the referral note and previous entries in the patient's hospital journal.

Hospitals in Norway also have National Priority Guidance Documents (NPGD) (24) they are required to follow when reviewing referrals. The purpose of the NPGDs is to ensure that the patient's right to health services are treated equal independently of geographical location and independent of the disease of the patient. All patient referrals are to be evaluated individually at the discretion of the senior physicians locally.

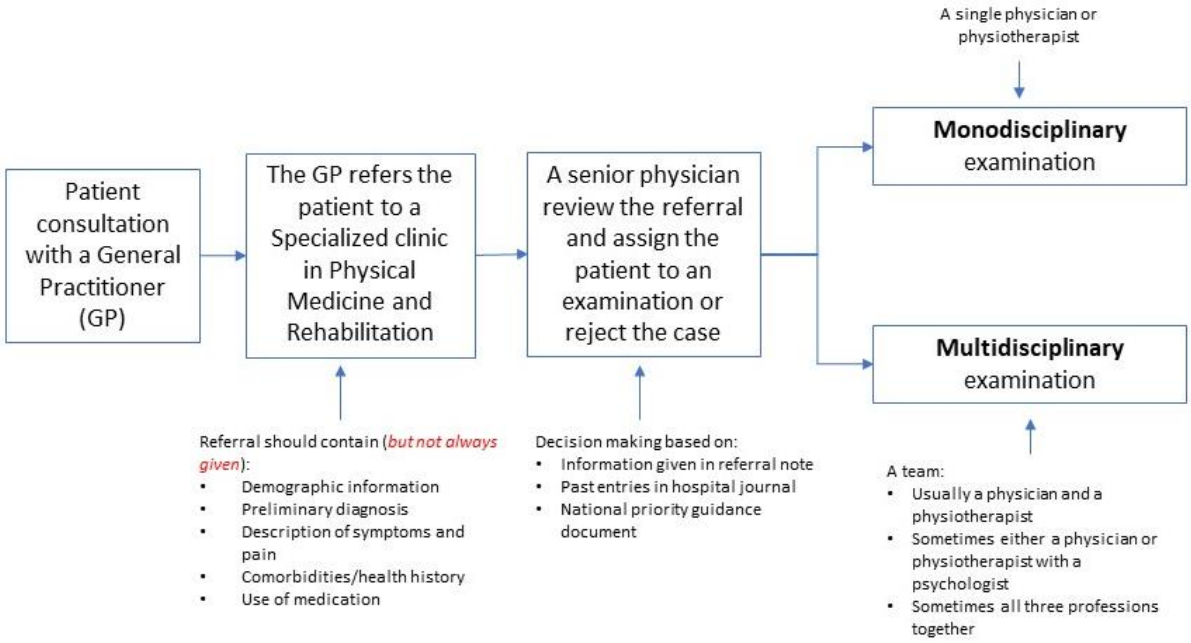
The NPGD for Physical medicine and Rehabilitation (25) have four chapters concerning neck pain and back pain. The NPGD state that loss of function, high level of pain and "yellow flags" can affect the patient's right to health services and length of waiting time. In these chapters there are no specific guidelines on which patients should be assigned to specific types of interventions.

At the specialized clinics in Physical medicine and rehabilitation at University Hospital in Northern Norway (UNN) back pain patients are assigned by a senior physician to either a

mono- or a multidisciplinary examination (Figure 1). Monodisciplinary examinations are carried out by one of the health professionals working at the clinic – commonly either a physician or a physiotherapist. Multidisciplinary examinations are carried out by a team of several health professionals – most commonly a physiotherapist and a physician.

At UNN the first consultation with the patient lasts about 1,5 hours. All examinations are single consultations that consist of a thorough and educational clinical examination, information about diagnosis and advice on self-care customized to the patient. The examinations by themselves are therefore regarded as an intervention. However, if there are findings suggesting that other examinations and/or interventions are necessary, the patient is referred further. UNN is connected to the Norwegian Neck and Back Registry.

Figure 1: Patient course at the Specialized clinic for Physical Medicine and Rehabilitation at UNN Tromsø



1.3 The Norwegian Neck and Back Registry

"Norsk nakke- og ryggregister" or the Norwegian Neck and Back Registry (NNRR) is a medical quality registry which collected data from 4 of the 15 specialized clinics in physical medicine and rehabilitation at the hospitals in Norway in 2018 (26). Data for the registry is collected through several self-report patient questionnaires and a health professional questionnaire. The questionnaires collect demographic information and Patient Reported Outcome Measures (PROMs) concerning physical and mental health, function and quality of life, as well as registering information concerning diagnosis and treatment recommendations.

The purpose of the NNRR is to improve on the quality of patient services at the specialized clinics. NNRR state that the registry will achieve this through (26):

- Evaluate geographical differences between the specialized clinics
- Evaluate which patients are referred to the specialized clinics
- Aid the process of assigning the right patients to the right treatments or interventions.
- Evaluate the medical and public health effect of the patient interventions
- Improve quality locally and eventually nationally.

NNRR is one of 51 national medical quality registers in Norway. The purpose of having national quality registries are documentation of effect of different treatments, as well as being a source for quality improvements and research (27).

Norway's four major specialized clinics in physical medicine and rehabilitation in Oslo, Bergen, Trondheim and Tromsø all have mono- and multidisciplinary examinations as the first consultation of back pain patients while registering patient data in the Norwegian Neck and Back Registry (NNRR). There is great variation in how many patients the clinics assign to either type of examinations. According to the NNRR annual report for 2018 (28) multidisciplinary examination is reported in 99.6% of all examinations in Bergen. Comparatively, in Trondheim only 2.5% of the patients had a multidisciplinary examination. While in Tromsø, multidisciplinary examinations were registered for 11.1% of the patients.

1.4 Significance of the investigation

Multidisciplinary examinations require more resources as more than one health professional are present. It is therefore important to investigate if the patient groups selected to mono- or multidisciplinary examinations differ with today's practice. In addition, additional knowledge of the patient characteristics can be beneficial in the tailoring of targeted interventions for the patients. It can also shed light on one of the components of multidisciplinary interventions. Many studies have investigated interventions and risk factors of back pain, but to our knowledge we currently have scare documentation regarding the characteristics of the patients assigned to either mono- or multidisciplinary examinations in a specialized clinic for physical medicine and rehabilitation.

The hypothesis would be that "more complex" patients are assigned to multidisciplinary examinations, but no studies have confirmed this hypothesis. More complex patients would be patients with more yellow flags, worse pain, worse function and potentially requiring different treatments.

1.5 Research question and aims

The purpose of this study was to investigate differences in neck- and/or back-pain patients either assigned to mono- or multidisciplinary examinations at UNN in 2018 who filled out the NNRR first patient questionnaire.

The specific aims include:

- To describe patient characteristics
- To investigate if there are differences between mono- and multidisciplinary examination concerning yellow flags, pain indicators and function level
- To investigate if there are differences between mono- and multidisciplinary examination concerning further treatment recommendations

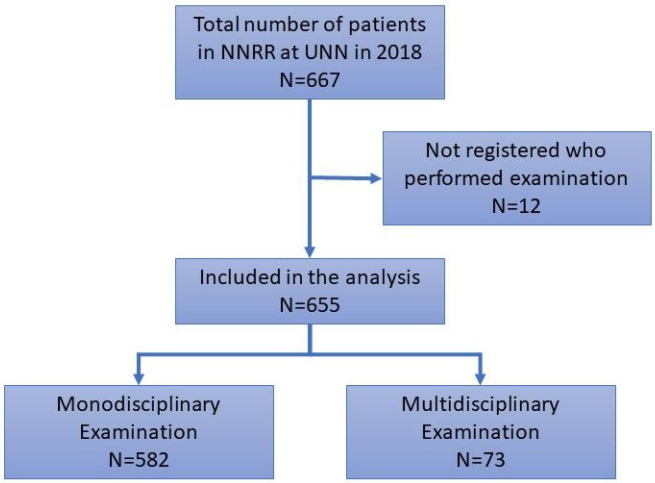
2 Material and method

2.1 Study population

The study population consist of all patients referred to the specialized clinic in physical medicine at UNN and who were eligible for inclusion in the NNRR in 2018. 87% of the eligible patients completed the first patient questionnaire of the NNRR (28).

The original dataset contained information of all the 667 patients who were examined at UNN and who completed the NNRR first patient questionnaire during 2018. All patients with no registration of which health professional(s) performed the examination were excluded (N=12). The final study sample therefore contained 655 patients.

Figure 2: Study Overview



2.2 Information on NNRR questionnaires

The NNRR collect data through a total of four questionnaires. The data used in this analysis are sourced from the two NNRR questionnaires used at the patient's first consultation at the hospital; "the first patient questionnaire" and "the health professional questionnaire".

The first patient questionnaire is a self-report questionnaire filled out electronically by a patient in the waiting room just prior to the examination. Therefore, it is filled out after the patient has been assigned to either a mono- or multidisciplinary examination. Sex, age, family status, education level, smoking status, physical activity level, pain duration, pain intensity, pain location, pain medication use, pain beliefs, work satisfaction, feeling of not being wanted

back at work, feeling of having a physical demanding or monotonous work, past surgery and the patient-reported outcome measures (PROMs) are sourced from the first patient questionnaire. The health professional(s) have access to the information gathered through the first patient questionnaire during the examination and it is used as a clinical tool.

The health professional questionnaire is filled out right after the examination by the health professional(s). The health professional(s) use information gathered in the examination and from the patient's hospital record to fill out the questionnaire. In this study information on which health professional(s) performed the examination, the patient's employment and sick leave status, as well as the diagnosis and treatment recommendations given during the examination are sourced from the health professional questionnaire.

There were also several other variables available from both questionnaires in the original dataset which were left out, as they were not relevant for the aims of this study. Radiological findings were available from the health professional questionnaire in the dataset and arguably relevant for the aims of this study. Though it is not known when this information becomes available in the patient course (Figure 1). It could be made available both before and after the patients are assigned to an examination. This leads to a potential selection bias which is not possible to account for when trying to analyze the patient's radiological findings in the scope of this study. Preliminary sensitivity analysis was performed though, and no statistically significant differences were detected in the available radiological variables and the patients assigned to mono- or multidisciplinary examinations. These results are available in the appendix.

2.3 Information on multidisciplinary classification

The health professional questionnaire registers which health professional(s) met the patient during the examination in four separate dichotomous variables; Doctor (yes, no), Nurse (yes, no), Physiotherapist (yes, no) and Other (yes, no). This data was used to compute a new dichotomous variable; "Multidisciplinary examination" (Yes/No). The category "No" is defined as only one health professional, while "Yes" is defined as two or more health professionals present under examination. These two groups have been compared according to the aims of the study.

Throughout this text the patients in the category “Multidisciplinary examination: Yes” are referred to as the “Multidisciplinary group”. The patients in the category “Multidisciplinary examination: No” are referred to as the “Monodisciplinary group”.

2.4 Information on patient characteristics

Patient characteristics reported were age (continuously), sex (male, female), family status (single, partner, married), education level (primary school, vocational subjects, secondary school, <4 years college or university, >4 years college or university), smoking status (yes, no), physical activity level (sedentary, light, moderate, hard).

2.5 Information on work-related variables

Work-related variables reported were unemployed (yes, no), on current sick leave (yes, no), on disability pension (yes, no), felt wanted back by employer (yes, no), previous sick leave (0 times, 1 time, 2-5 times, 6-10 times, >10 times), work satisfaction (continuous from worst to best), feeling of having a physical demanding job (continuous from not at all to very demanding) and feeling of having a monotonous job (continuous from not at all to very monotonous). Continuous variables were measured using Visual Analogue Scales (VAS) 0-10.

2.6 Information on diagnosis

Diagnosis are set by the health professional(s) during the examination and registered in the health professional questionnaire.

In the dataset there were a total of 60 different ICD-10 diagnosis codes reported. The patients could have registered up to six separate diagnosis. All diagnosis was recoded from the 60 different ICD-10 codes listed in the dataset into one variable (“diagnosis”) with three categories; “non-specific LBP”, “non-specific neck pain” and “other or multiple diagnosis” (see appendix for ICD-10 code list with recoding).

“Non-specific LBP” include patients who have received the ICD-10 diagnosis “M54.5 Lumbago” and no other additional diagnosis. “Non-specific neck pain” include patients who have received the ICD-10 diagnosis “M54.2 Neck pain” and no other additional diagnosis.

The “other or multiple diagnosis” contain patients with all other diagnosis or patients with multiple diagnosis. The patients with multiple diagnosis also include patients with “M54.5 Lumbago” or “M54.2 Neck pain” who also had one or more additional diagnosis.

The recoding was done to be able to estimate which types of diagnosis received multi- or monodisciplinary examinations. The categories were chosen based on “non-specific LBP” and “non-specific neck pain” being the two largest identifiable uniform subgroups among the diagnosis codes.

An important note on diagnosis; a patient’s potential mental health problems are not subjected to be diagnosed at the specialized clinic for physical medicine and rehabilitation. If any mental health problems were detected during the examination, it would only be described in the written text of the hospital record. Therefore, we have no information on potential diagnosis related to mental health that a patient could have received prior or after the examination.

2.7 Information on pain indicators, pain beliefs, pain medication and past surgery

Pain indicators reported in the dataset were pain intensity in activity (continuous), pain intensity in rest (continuous), pain duration (no pain, <3 months, 3-12 months, 1-2 years, >2 years) and number of pain regions (continuous).

Pain beliefs are reported through two variables; pain is caused by work (yes, no) and pain is caused by mental problems (yes, no).

Pain medication use is reported through two variables; non-prescription pain medication use (not in the last month, less than every week, more than every week while not daily, daily) and prescription pain medication use (not in the last month, less than every week, more than every week while not daily, daily).

Previous surgeries are reported through two variables; back surgery (yes, no) and neck surgery (yes, no).

Both pain intensity in rest and activity are reported by the patients through a VAS 0-10 with 0 being no pain and 10 being worst imaginable pain. Mean score for both are reported for the mono- and multidisciplinary groups.

The patients originally complete a pain drawing with 32 regions in the first patient questionnaire. The dataset lists all painful regions of the pain drawing for each patient. The mean number of pain regions are reported for the mono- and multidisciplinary groups.

Pain duration was collapsed from five categories to four categories (<3 months, 3-12 months, 1-2 years, >2 years) and reported. This was due to very few patients in the original "no pain" category. The three patients reporting no pain were included in the <3 months category.

2.8 Information on PROMs

The following overview refers to different instruments used to assess Patient Reported Outcome Measures (PROMs) describing different types of functioning are reported through the first patient questionnaire:

Oswestry Disability Index

Oswestry Disability Index (ODI) (29) is a questionnaire that describes how back pain affect current daily functioning with 10 statements graded by the patient. A score is calculated from 0-100. Mean score are reported for the mono- and multidisciplinary group. The score was recoded into five categories described by the ODI; 0-20 = minimal disability, 21-40 = moderate disability, 41-60 = severe disability, 61-80 = crippled, 81-100 = bed-bound or exaggerating symptoms. Because of the low number in the "bed-bound or exaggerating symptoms" category it was collapsed into "crippled", and the remaining four categories were reported for the mono- and multidisciplinary group.

Neck Disability Index

Neck Disability Index (NDI) (30) describes how neck pain affect current daily functioning with 10 statements graded by the patient. A score is calculated from 0-50. Mean score are reported for the mono- and multidisciplinary group. The score was recoded into five categories as described by the NDI; 0-4 = no disability, 5-14 = mild disability, 15-24 = moderate disability, 25-34 = severe disability, 35-50 = complete disability. Because of the low numbers in "no disability" and "complete disability", the original five categories were collapsed to three categories (no/mild disability, moderate disability, severe/complete disability) and reported for the mono- and multidisciplinary groups.

Hopkins Symptoms Check List 10

Hopkins Symptoms Check List 10 (HSCL-10) (31, 32) is a questionnaire that aims to identify mental health problems last seven days through 10 items graded from 1 to 4 by the patient, with 4 being the highest severity. An individual score from 0 to 4 is calculated for each patient. Mean scores are reported for the mono- and multidisciplinary groups. A score higher

than 1,85 indicates mental distress and a new variable was calculated (HSCL-10 >1.85: yes, no) and reported for the mono- and multidisciplinary group.

Fear Avoidance Beliefs Questionnaire

Fear Avoidance Beliefs Questionnaire (FAB-Q) (33) registers the patients fear avoidance beliefs towards physical activity and work. The questionnaire provides four statements about physical activity and seven statements about work and the patient is asked to rate. Each statement is rated from 0 to 6 with 0 being in full disagreement with the statement and 6 being in full agreement. Higher scores indicate increased fear-avoidance beliefs. The two separate scores for physical activity (max 24 points) and work (max 42 points) and one combined score (max 66 points) are calculated for each patient. Mean scores of all three FAB-Q scores are reported for the mono- and multidisciplinary group.

Ursin Health Inventory

Ursin Health Inventory (UHI) (34) is a questionnaire that registers 29 (33 for women) different single items of subjective health complaints (SHCs) experienced last 4 weeks. The single items are graded in severity by the patient from 0 to 3 (0 = none, 1 = some, 2 = much, 3 = severe). It is suggested in the presentation of the UHI scoring system that reporting frequencies of single items can be the most useful, and with a cutoff of > 0. It is also suggested that for specific groups the cutoff can be elevated to > 1 or > 2. It was decided to use a cutoff of > 1 in this study. Each of the 33 single items were recoded from the original four categories to two categories (0 = 0 and 1, 1 = 2 and 3). According to the aims of this study the frequencies of selected recoded single items was reported. The single items which are reported: headache (0, 1), upper back pain (0, 1), shoulder pain (0, 1), arm pain (0, 1), anxiety (0, 1), depression (0, 1), tiredness (0, 1), dizziness (0, 1), stomach discomfort (0, 1), stomach pain (0, 1), gas discomfort (0, 1) and obstipation (0, 1).

EQ-5D-3L[1]

EQ-5D-3L (35) is a descriptive system unrelated to diagnosis meant to create a score which indicates the patient's health-related quality of life. Measures 5 different dimensions of current living on a 3-level scale. A separate Visual Analogue Scale (0 to 100 with 0 being

[1] The name «EQ-5D-3L» is not an abbreviation, but the name which the user guide state as the correct term to use in print or verbally.

"worst imaginable health" and 100 being "best imaginable health") is also provided called EQ-VAS. Mean scores for both EQ-5D-3L and EQ-VAS are reported for the mono- and multidisciplinary groups.

2.9 Information on treatment recommendations

The health professional(s) list all recommended treatments in the primary care; physician (yes, no), physiotherapist (yes, no), manual therapist (yes, no), chiropractor (yes, no), psychologist (yes, no) and work related follow-up (yes, no).

The health professional(s) also list all recommended treatments in the specialist services; treated in any specialist services (yes, no), treated in own specialist service (yes, no), referred to rehabilitation center (yes, no), referred to evaluation by other specialist (yes, no), referred to potential operation (yes, no), control after examination or treatment (yes, no), individual follow-up 1-2 times (yes, no), group treatment (none, 1-3 days, 4-10 days, >10 days).

In addition, the health professional(s) register several other recommendations regarding treatment of the patient; training activation (yes, no), work related follow-up (yes, no), cognitive approach (yes, no), education (yes, no) and psychomotoric physiotherapy (yes, no).

If the patient received no treatment recommendations it was registered in "no treatment (yes, no)".

At UNN there are only two group treatments offered in own specialist service; 2 days- and 10 days group treatment. The four original categories (none, 1-3 days, 4-10 days, >10 days) were recoded to fit the group treatments offered at UNN and two new variables were created; 2 days group treatment (yes = 1-3 days, no = none, 4-10 days, >10days) and 10 days group treatment (yes = 4-10 days and >10 days, no = none and 1-3 days) and reported for the mono- and multidisciplinary group.

A third new variable - any group treatment - was also calculated by summarizing the two new variables (2 days group treatment, 10 days group treatment) and reported for the mono- and multidisciplinary group. This was done to be able to investigate if there were differences in group treatments recommended overall for the mono- and multidisciplinary group.

2.10 Statistical analysis

This is an observational study with a cross-sectional design.

The descriptive statistics of the patients are presented by crude numbers, percentages, means, standard deviations and number of missing cases. The Chi-Square test was used to test for differences in all categorical variables between the mono- and multidisciplinary group. The Independent Samples T-test was used to compare differences in means of the continuous variables between the mono- and multidisciplinary group. P-values of crude numbers were calculated and presented. Missing information was identified, reported and excluded from further analysis.

Binary logistical regression analysis was used to calculate multivariable adjusted odds ratios (ORs) with 95% confidence intervals (95% CI) of the associations between mono- and multidisciplinary examinations and the various variables describing yellow flags, pain, function and treatment recommendations.

“Age”, “Sex” and “Smoking status” were considered as possible confounders due to association with several of the listed exposure variables and mono- and multidisciplinary examinations. However, just age was associated with both exposure variables and the outcome. Therefore, age was the only confounder adjusted for in the binary logistical regression analysis.

All analysis was done in IBM SPSS Statistics for Windows, Version 26.

2.11 Ethics

Permissions

A formal inquiry was sent to the Regional committee for medical- and health-related research (REK) who replied that this study does not require REK-approval. REK regarded this study as a local quality assessment project which only required local approval from the patient safety officer at UNN and approval from the NNRR board of directors.

An application to the UNN patient safety officer to use the data registered in the NNRR first patient and health professional questionnaire at UNN in 2018 was sent, as well as a formal request to the NNRR board of directors. Permission to use the data requested was granted by both parties.

Privacy and confidentiality

The manager of the NNRR merged the data in the first patient questionnaire and the data in the health professional questionnaire into one dataset to be used in this study. The individual patients in the dataset were “unidentified” by the manager of the NNRR before the dataset was handed over to the research group.

Informed consent process

All patients who complete the first patient questionnaire have given written consent allowing the NNRR data to be stored and used in research.

3 Results

3.1 Multidisciplinary classification

Of a total of 655 patients, 572 underwent a monodisciplinary examination, and were classified as the monodisciplinary group. The remaining 73 patients underwent a multidisciplinary examination and were classified as the multidisciplinary group.

3.2 Descriptive statistics

Table 1 to table 5 contain crude frequencies and mean scores from the first patient questionnaire and the health professional questionnaire. In dichotomous variables only answers with a value of 1 or yes has been listed unless other noted in the tables. This was done to limit the size of the tables.

3.2.1 Patient characteristics

There was a statistically significant mean difference of 5,6 years in age between the multi- and monodisciplinary group. The multidisciplinary group had a mean age of 39.3 (range = 17-65), while the monodisciplinary group had a mean age of 44.9 (range = 17-87) (Table 1).

44.4% had a family status as single in the multidisciplinary group compared to 28.1% in the monodisciplinary group (Table 1). This difference in family status was statistically significant.

19.5% in the monodisciplinary group were smokers compared 11.1% in the multidisciplinary group (**Feil! Fant ikke referansekilden.**), however this difference was not statistical significant.

3.2.2 Work-related variables

A higher proportion had previous sick leave(s) in the multidisciplinary group compared to the monodisciplinary group (Table 1), but this difference was not statistically significant.

In the work-related questions there were a higher number of non-response than in the other variables (Table 1). This was mainly related to patients who didn't have current employment. There was notably an even higher rate of non-response for the variable "not felt wanted back by employer", though no reason for this could be identified.

Table 1: Differences in patient characteristics and work-related variables between mono- and multidisciplinary examinations among back pain patients from the University Hospital in Northern Norway

	Examination type		p-value[3]	Missing
	Mono-disciplinary	Multi-disciplinary		
Sex				
Female, N (%)	316 (54.3)	40 (55.6)	0.936[1]	1
Male, N (%)	266 (45.7)	32 (44.4)		
Age				
Mean years (SD) (Range)	44.9 (13.9) (17-87)	39.3 (13.4) (17-65)	0.001[2]	
Family status				
Married, N (%)	254 (44.1)	25 (34.7)	0.017[1]	7
Partner, N (%)	160 (27.8)	15 (20.8)		
Single, N (%)	162 (28.1)	32 (44.4)		
Education				
Primary school, N (%)	88 (15.3)	12 (16.7)	0.274	7
Vocational subjects, N (%)	213 (37.0)	27 (37.5)		
Secondary school, N (%)	67 (11.6)	14 (19.4)		
< 4 years college or university, N (%)	119 (20.7)	12 (16.7)		
> 4 years college or university, N (%)	89 (15.5)	7 (9.7)		
Smoking				
Smoker, N (%)	111 (19.5)	8 (11.1)	0.053[1]	14
Physical activity level				
Sedentary	90 (15.9)	11 (15.1)	0.959[1]	15
Light	363 (64.0)	49 (67.1)		
Moderate	89 (15.7)	10 (13.7)		
Hard	25 (4.4)	3 (4.1)		
Work-related variables				
Unemployed, N (%)	17 (2.9)	3 (4.1)	0.389[1]	0
On sick leave, N (%)	209 (35.9)	29 (39.7)	0.523[1]	0
Previous sick leave, N (%)			0.059[1]	45
No	312 (58.0)	38 (52.8)		
1 time	104 (19.3)	12 (16.7)		
2-5 times	103 (19.1)	17 (23.6)		
6-10 times	9 (1.7)	0 (0.0)		
>10 times	10 (1.9)	5 (6.9)		
Disability pension, N (%)	46 (7.9)	7 (9.6)	0.376[1]	0
Not felt wanted back by employer, N (%)	75 (16.0)	13 (21.7)	0.266[1]	126
Work satisfaction, mean VAS (SD)	7.44 (2.7)	7.05 (2.66)	0.273[2]	78
Feeling of having a physical demanding job, mean VAS (SD)	4.89 (3.17)	4.52 (2.97)	0.376[2]	63
Feeling of having a monotonous job, mean VAS (SD)	4.24 (2.91)	4.52 (2.97)	0.471[2]	76

[1] Chi-Square Test, [2] Independent samples t-test, [3] Fischer's exact p-value listed for variables with groups containing less than 10 patients

3.2.3 Diagnosis

No statistically significant differences between the multi- and monodisciplinary groups were detected across the diagnosis categories (Table 2).

Table 2: Differences in diagnosis classification between mono- and multidisciplinary examinations among back pain patients from the University Hospital in Northern Norway

	Examination type		p-value	Missing
	Mono-disciplinary	Multi-disciplinary		
Diagnosis				
Non-specific LBP, N (%)	267 (46.7)	31 (43.1)	0.832[(1)]	11
Non-specific neck pain, N (%)	78 (13.6)	11 (15.3)		
Other or multiple diagnosis, N (%)	227 (39.7)	30 (41.7)		

[1] Chi-Square Test

3.2.4 Pain indicators, pain beliefs, pain medication and past surgery

No statistically significant differences were registered among the different pain indicators between the mono- and multidisciplinary group (Table 3).

Among pain beliefs 16.4% of the patients in the multidisciplinary group believed the cause of pain were due to mental problems compared to 7.2% in the monodisciplinary group, and this difference was statistically significant.

There was a statistically significant difference in prescription pain medication use between the mono- and multidisciplinary group (Table 3). A higher proportion of the multidisciplinary group reported using prescription pain medication.

High numbers of non-response were registered in both non-prescription and prescription pain medication use (Table 3) and investigated further. It was found that 73 of the patients who had answered the question of prescription pain medication use left out answering non-prescription pain medication use. Of the patients who had answered the question of non-prescription pain medication use 69 left out answering the question of prescription pain medication use.

Table 3: Differences in pain indicators, pain medication usage and previous surgery between mono- and multidisciplinary examinations among back pain patients from the University Hospital in Northern Norway

	Examination type		p-value(3)	Missing
	Mono-disciplinary	Multi-disciplinary		
Pain indicators				
Pain intensity in rest, mean VAS (SD)	5.36 (2.3)	5.33 (2.1)	0.919(2)	6
Pain intensity in activity, mean VAS (SD)	6.51 (2.2)	6.78 (1.94)	0.307(2)	8
Number of regions marked on pain drawing, mean (SD)	6.86 (5.43)	7.26 (5.59)	0.556(2)	6
Pain duration:			0.303(1)	7
< 3 months	37 (6.4)	2 (2.7)		
3-12 months	145 (25.2)	24(32.9)		
1-2 years	86 (15.0)	8 (11.0)		
>2 years	307 (53.4)	39 (53.4)		
Pain beliefs				
Pain is caused by work, N (%)	230 (39.5)	31 (42.5)	0.628(1)	0
Pain is caused by mental problems, N (%)	42 (7.2)	12 (16.4)	0.007(1)	0
Use of pain medication				
Non-Prescription:			0.114(1)	82
Not in the last month	120 (23.6)	22 (34.4)		
< every week	79 (15.5)	5 (7.8)		
> every week	167 (32.8)	23 (35.9)		
Every day	143 (28.1)	14 (21.9)		
Prescription:			0.032(1)	86
Not in the last month	171 (34.1)	16 (23.5)		
< every week	63 (12.6)	5 (7.4)		
> every week	111 (22.2)	25 (36.8)		
Every day	156 (31.1)	22 (32.4)		
Past surgery				
Back surgery, N (%)	63 (11.0)	10 (13.7)	0.452(1)	15
Neck surgery, N (%)	17 (3.0)	2 (2,7)	0.631(1)	13

1 Chi-Square Test, 2 Independent samples t-test, 3 Fischer's exact p-value listed for variables with groups containing less than 10 patients

3.2.5 PROMs

66.2% of the patients in the multidisciplinary group had a HSCL-10 Score >1.85 compared to 51.3% in the monodisciplinary group, and this difference was statistically significant. The mean HSCL-10 Score was also statistically significantly higher in the multidisciplinary group compared to the monodisciplinary group (Table 4)

Among the UHI single items there were several statistically significant differences between the groups (Table 4). 22.5% of the patients in the multidisciplinary group reported depression compared to 11.7% in the monodisciplinary group. Anxiety was reported by 17.1% in the multidisciplinary group compared to 9.2% in the monodisciplinary group. In addition, statistically significant higher proportions of the multidisciplinary group compared to the monodisciplinary group reported stomach discomfort, upper back pain and shoulder pain.

No other statistically significant differences between the mono- and multidisciplinary group were detected in the results from the PROMs. However, the multidisciplinary group reported a higher mean NDI score compared to the monodisciplinary group (Table 4). In addition, in the NDI categories 46.9% reported moderate disability and 24.5% reported severe/complete disability in the multidisciplinary group, compared to 41.9% and 19.4% in the monodisciplinary group.

In addition, table 4 shows the numbers of nonresponse. The number of nonresponses in the work score of the FAB-Q was mainly due to patients being not employed at the time. Only patients with both FAB-Q physical activity score and work score were reported in the total score. 156 patients were non-responders in the EQ-VAS score.

Table 4: Differences in PROMs between mono- and multidisciplinary examinations among back pain patients from the University Hospital in Northern Norway

	Examination type		p-value	Missing
	Mono-disciplinary	Multi-disciplinary		
ODI				
Score, mean (SD)	30.75 (14.01)	30.19 (11.94)	$0.744_{(2)}$	4
Categorical, N (%)			$0.795_{(1)}$	4
1. Minimal disability	161 (27.9)	20 (27.4)		
2. Moderate disability	290 (50.2)	40 (54.8)		
3. Severe disability	111 (19.2)	12 (16.4)		
4. Crippled	16 (2.8)	1 (1.4)		
NDI				
Score, mean (SD)	17.34 (8.55)	19.22 (7.35)	$0.144_{(2)}$	0[2]
Categorical			$0.369_{(1)}$	0[3]
1. None/mild disability	132 (38.7)	14 (28.6)		
2. Moderate disability	143 (41.9)	23 (46.9)		
3. Severe/complete disability	66 (19.4)	12 (24.5)		
HSCL-10				
Score, mean (SD)	1.98 (0.66)	2.15 (0.62)	$0.041_{(2)}$	23
Score > 1,85, N (%)	288 (51.3)	47 (66.2)	$0.018_{(1)}$	23
FAB-Q				
Total Score, mean (SD)	32.37 (15.12)	31.46 (14.92)	$0.643_{(2)}$	58
Physical activity Score, mean (SD)	11.82 (6.04)	12.22 (5.89)	$0.592_{(2)}$	23
Work Score, mean (SD)	20.45 (11.66)	19.57 (11.75)	$0.562_{(2)}$	56
UHI Single items				
Headache	240 (42.9)	29 (42.9)	$0.964_{(2)}$	28
Upper back pain	232 (42.6)	40 (58.0)	$0.015_{(2)}$	41
Shoulder pain	247 (44.6)	42 (58.3)	$0.028_{(2)}$	29
Arm pain	187 (33.7)	26 (37.7)	$0.510_{(2)}$	31
Anxiety	51 (9.2)	12 (17.1)	$0.036_{(2)}$	28
Depression	65 (11.7)	16 (22.5)	$0.011_{(2)}$	29
Stomach discomfort	55 (10.0)	13 (18.3)	$0.035_{(2)}$	34
Stomach pain	61 (11.2)	12 (17.1)	$0.151_{(2)}$	42
Gas discomfort	105 (19.0)	12 (16.7)	$0.635_{(2)}$	30
Obstipation	46 (8.3)	10 (13.9)	$0.116_{(2)}$	27
Dizziness	99 (17.9)	11 (15.3)	$0.582_{(2)}$	30
Tiredness	305 (54.2)	33 (46.5)	$0.221_{(2)}$	21
EQ-5D-3L				
Score, mean (SD)	0.475 (0.32)	0.480 (0.31)	$0.906_{(2)}$	32
EQ-VAS, mean VAS (SD)	53.14 (19.06)	52.78 (19.70)	$0.896_{(2)}$	156

[1] Chi-Square Test, [2] Independent samples t-test, PROMs = Patient Reported Outcome Measures, ODI = Oswestry Disability Index, NDI = Neck Disability Index, HSCL-10 = Hopkins Symptoms Checklist 10, FAB-Q = Fear Avoidance Beliefs Questionnaire, UHI = Ursin Health Index

[2] Only patients with neck pain are requested to report the NDI. 265 patients were not listed with an NDI score. It is not possible to detect an accurate number of missing for this variable, as there is no possibility to separate patients who did not have neck pain from patients with neck pain choosing to be non-responders. Though it is very likely that a true number of non-responders is small.

[3] Same as [2].

3.2.6 Treatment recommendations

There were several differences between the mono- and multidisciplinary group in the treatment recommendations.

38.0% of the patients in the monodisciplinary group received no treatment recommendations compared to 23.3% in the multidisciplinary group (Table 5). This difference was statistically significant.

41.1% of the patients in the multidisciplinary group were recommended follow-up by a physician in primary care compared to 18.2% in the monodisciplinary group. This difference was statistically significant. 6.8% of the patients in the multidisciplinary group were recommended work-related follow-up in primary care compared to 2.2% in the monodisciplinary group (Table 5). This difference was also statistically significant.

A larger proportion of the patients in the multidisciplinary group were also recommended to follow-up by a physiotherapist in the primary care, however this difference was not statistically significant (Table 5).

There were no statistically significant differences in recommendations among treatments or referrals in the specialist services between the mono- and multidisciplinary group (Table 5).

Among other recommendations there were differences in training activation and work-related follow-up between the mono- and multidisciplinary group. A larger portion of the patients in the monodisciplinary group received training activation recommendation, while a larger portion of the multidisciplinary group received work-related follow-up recommendation (Table 5). These differences were however not statistically significant.

Table 5: Differences in treatment recommendations between mono- and multidisciplinary examinations among back pain patients from the University Hospital in Northern Norway

	Examination type		p-value[(2)]
	Mono-disciplinary	Multi-disciplinary	
No treatment, N (%)	226 (38.8)	17 (23.3)	0.010[(1)]
In primary care			
Physician, N (%)	106 (18.2)	30 (41.1)	<0.001[(1)]
Physiotherapist, N (%)	248 (42.6)	36 (49.3)	0.276[(1)]
Manual therapist, N (%)	25 (4.3)	3 (4.1)	0.619[(1)]
Chiropractor, N (%)	9 (1.5)	1 (1.4)	0.692[(1)]
Psychologist, N (%)	10 (1.7)	2 (2.7)	0.394[(1)]
Work related follow-up, N (%)	13 (2.2)	5 (6.8)	0.040[(1)]
In specialist services			
Treated in any specialist services, N (%)	319 (54.8)	36 (49.3)	0.374[(1)]
Treated in own specialist service, N (%)	142 (24.4)	17 (23.3)	0.835[(1)]
Referred to rehabilitation center, N (%)	71 (12.2)	10 (13.7)	0.415[(1)]
Referred to evaluation by other specialist, N (%)	26 (4.5)	2 (2.7)	0.377[(1)]
Referred to potential operation, N (%)	26 (4.5)	3 (4.1)	0.592[(1)]
Control after examination or treatment, N (%)	78 (13.4)	8 (11.0)	0.357[(1)]
Individual follow-up 1-2 times, N (%)	33 (5.7)	4 (5.5)	0.603[(1)]
Any group treatment, N (%)	122 (21.0)	15 (20.5)	0.935[(1)]
2 days group treatment, N (%)	52 (8.9)	7 (9.6)	0.494[(1)]
10 days group treatment, N (%)	70 (12.0)	8 (11.0)	0.486[(1)]
Other recommendations			
Training activation, N (%)	239 (41.1)	22 (30.1)	0.072[(1)]
Work related follow-up, N (%)	68 (11.7)	14 (19.2)	0.068[(1)]
Cognitive approach, N (%)	144 (24.7)	17 (23.3)	0.786[(1)]
Education, N (%)	133 (22.9)	17 (23.3)	0.933[(1)]
Psychomotoric physiotherapy, N (%)	23 (4.0)	1 (1.4)	0.230[(1)]

[1] Chi-Square Test, [2] Fischer's exact p-value listed for variables with groups containing less than 10 patients

3.3 Associations between variables describing yellow flags and mono- and multidisciplinary examinations

Among the yellow flags describing mental health problems there were several statistically significant associations between the multidisciplinary and monodisciplinary group (Table 6).

There was increased odds (OR=2.53, 95% CI=1.25-5.10) that the patient believed his/her pain was caused my mental problems in the multidisciplinary group compared to in the monodisciplinary group (Table 6). Among patients in the multidisciplinary group there were increased odds of reporting both depression (OR=2.04, 95% CI=1.09-3.80) and anxiety (2.06, 95% CI=1.03-4.12) compared to patients in the monodisciplinary group.

There were increased odds (OR=1.64, 95% CI=0.96-2.77) of having a HSCL-10 score above 1.85 in the multidisciplinary group compared to in the monodisciplinary group (Table 6). This variable was statistically significant in the crude numbers; however, it was no longer statistically significant when adjusted for age.

For comorbidities there were increased odds of stomach discomfort (OR=2.16, 95% CI=1.10-4.24), upper back pain (OR=1.74 95% CI=1.04-2.90) and shoulder pain (OR=1.85 95% CI=1.12-3.06) in the multidisciplinary group compared to the monodisciplinary group (Table 6).

There were also increased odds (OR=4.54, 95% CI=1.42-13.95) of the patient having been on sick leave more than 10 times previously in the multidisciplinary group compared to the monodisciplinary group (Table 6).

Table 6: Odds ratios (OR) and 95% confidence intervals (CI) for the association between yellow flags and being assigned to multidisciplinary examination among back pain patients from the University Hospital in Northern Norway

	Odds Ratio	95% CI	p-value
Work related issues and sick leave			
Unemployed	1.080	0.303-3.853	0.906
On sick leave	1.089	0.659-1.802	0.739
Previous sick leave (ref: no)			
1 time	0.844	0.422-1.687	0.631
2-5 times	1.255	0.675-2.332	0.473
6-10 times [1]	-	-	-
>10 times	4.454	1.422-13.946	0.010
Not felt wanted back by employer	1.460	0.749-2.847	0.266
Work satisfaction (one-point difference)	0.966	0.877-1.063	0.476
Pain is caused by work	1.083	0.658-1.780	0.754
Mental health problems			
Pain is caused by mental problems	2.527	1.253-5.096	0.010
HSCL-10 Score >1.85	1.635	0.964-2.774	0.068
HSCL-10 Score	1.294	0.914-1.833	0.146
UHI Single items			
Anxiety	2.058	1.027-4.123	0.042
Depression	2.038	1.093-3.799	0.025
Comorbidities			
Number of regions marked on pain drawing	1.012	0.970-1.057	0.572
Diagnosis (ref: Non-specific LBP)			
Non-specific neck pain	1.399	0.665-2.946	0.376
Other or multiple diagnosis	1.277	0.743-2.193	0.376
UHI Single items			
Stomach discomfort	2.162	1.102-4.241	0.025
Stomach pain	1.517	0.766-3.003	0.232
Gas discomfort	0.880	0.455-1.702	0.703
Obstipation	1.814	0.863-3.809	0.116
Tiredness	0.696	0.422-1.148	0.696
Dizziness	0.816	0.412-1.617	0.561
Headache	0.874	0.521-1.468	0.611
Upper back pain	1.738	1.041-2.900	0.034
Shoulder pain	1.849	1.117-3.060	0.017
Arm pain	1.362	0.802-2.315	0.253
Past debilitating back pain			
Back surgery	1.616	0.773-3.378	0.202
Neck surgery	1.717	0.534-5.520	0.364
Pessimistic or negative attitude towards the pain			
FAB-Q Total Score	0.994	0.977-1.011	0.469
FAB-Q Physical Activity Score	1.006	0.964-1.049	0.786
FAB-Q Work Score	0.991	0.969-1.013	0.410

Binary Logistic Regression with 'Multidisciplinary examination (Yes, No)' as dependent variable
Adjusted for: Age
[1] There were no patients in the 6-10 times category in the multidisciplinary group, and therefore unable to calculated odds ratio for this category.

3.4 Associations between variables describing pain and function and mono- and multidisciplinary examinations

Two statistically significant associations concerning pain medication use between the mono- and multidisciplinary group were detected. There were less odds (OR=0.34, 95% CI=0.12-0.94) of patients in the multidisciplinary group using non-prescription pain medication less than every week compared to the patients in the monodisciplinary group. In addition, there were increased odds (OR=2.36, 95% CI=1.20-4.64) of the patients in the multidisciplinary group using prescription pain medication more than once every week compared to the patients in the monodisciplinary group (Table 7).

No other statistically significant associations were detected between the remaining variables for pain and function and the mono- and multidisciplinary group.

However, there were increased odds (OR=1.72, 95% CI=0.86-3.43) of the patients in the multidisciplinary group using prescription pain medication every day compared to the monodisciplinary group (Table 7).

In addition, there were increased odds (OR=1.66 95% CI=0.81-3.41) for being in the NDI category of moderate disability and increased odds (OR=1.96, 95% CI=0.84-4.55) for being in the NDI category of severe/complete disability for patients in the multidisciplinary group compared to patients in the monodisciplinary group (Table 7).

Table 7: Odds ratios (OR) and 95% confidence intervals (CI) for the association between variables describing pain and function and being assigned to multidisciplinary examination among back pain patients from the University Hospital in Northern Norway

	Odds Ratio	95% CI	p-value
Pain intensity and duration			
Pain in rest	1.013	0.906-1.131	0.826
Pain in activity	1.074	0.955-1.208	0.230
Pain duration (ref: < 3 months)			
3-12 months	3.350	0.752-14.929	0.113
1-2 years	1.932	0.388-9.615	0.421
>2 years	2.839	0.652-12.366	0.165
Pain medication use			
Non-Prescription: (ref: Not in the last month)			
< every week	0.341	0.123-0.943	0.038
> every week	0.718	0.380-1.356	0.307
Every day	0.580	0.282-1.193	0.139
Prescription: (ref: Not in the last month)			
< every week	0.865	0.303-2.472	0.786
> every week	2.360	1.200-4.642	0.013
Every day	1.717	0.860-3.428	0.125
Function level			
ODI Score	1.000	0.982-1.018	0.972
ODI Categorical (ref: Minimal disability)			
Moderate disability	1.258	0.704-2.248	0.438
Severe disability	0.974	0.454-2.092	0.947
Crippled	0.676	0.079-5.173	0.676
NDI Score	1.031	0.995-1.069	0.089
NDI Categorical (Ref: None/mild disability)			
Moderate disability	1.664	0.812-3.409	0.164
Severe/Complete disability	1.956	0.842-4.545	0.119
EQ-5D-3L Score	1.104	0.497-2.452	0.808
EQ-VAS	1.000	0.986-1.015	0.963

Binary Logistic Regression with 'Multidisciplinary examination (Yes, No)' as dependent variable
Adjusted for: Age

3.5 Associations between treatment recommendations and mono- and multidisciplinary examinations

Among the treatment recommendations there were several statistically significant associations between the multi- and monodisciplinary group:

There were less odds (OR=0.43, 95% CI =0.24-0.76) of the patients in the multidisciplinary group of being recommended no treatment at all compared to the patients in the monodisciplinary group (Table 8).

There were increased odds (OR=3.1, 95% CI=1.85-5.20) of being recommended follow-up by a physician in primary care in the multidisciplinary group compared to in the monodisciplinary group (Table 8).

There were increased odds (OR=3.56, 95% CI=1.22-10.45) of being recommended work-related follow-up in primary care among patients in the multidisciplinary group compared to in the monodisciplinary group (Table 8).

There were less odds (OR=0.54, 95% CI=0.32-0.92) of being recommended training activation in the multidisciplinary group compared to in the monodisciplinary group (Table 8). The training activation variable was not a statistically significant difference in the crude numbers between the mono- and multidisciplinary groups, but the variable was statistically significant when adjusted for age (Table 8).

Table 8: Odds ratios (OR) and 95% confidence intervals (CI) for the association between treatment recommendations and being assigned to multidisciplinary examination among back pain patients from the University Hospital in Northern Norway

	Odds Ratio	95% CI	p-value
No treatment recommendations	0.428	0.241-0.761	0.004
Treatment in the primary care			
Physician	3.099	1.847-5.199	<0.001
Physiotherapist	1.409	0.861-2.308	0.173
Manual Therapist	0.917	0.268-3.134	0.890
Chiropractor	0.870	0.107-7.074	0.896
Psychologist	1.451	0.308-6.828	0.638
Work-related follow-up	3.564	1.215-10.452	0.021
Treatment in specialist services			
Treated in any specialist services	0.771	0.472-1.261	0.301
Treated in own specialist service	0.875	0.490-1.563	0.652
Referred to rehabilitation center	1.288	0.625-2.652	0.493
Referred to evaluation by other specialist	0.627	0.144-2.722	0.533
Referred to potential operation	1.056	0.307-3.632	0.931
Control after examination or treatment	0.775	0.356-1.687	0.521
Individual follow-up 1-2 times	0.904	0.308-2.651	0.854
Any group treatment	0.882	0.480-1.620	0.686
2 days group treatment	1.004	0.435-2.317	0.992
10 days group treatment	0.815	0.373-1.782	0.609
Other recommendations			
Training activation	0.538	0.315-0.920	0.023
Work related follow-up	1.641	0.864-3.115	0.130
Cognitive approach	0.792	0.441-1.420	0.433
Education	0.938	0.524-1.679	0.829
Psychomotoric physiotherapy	0.330	0.044-2.488	0.282

Binary Logistic Regression with 'Multidisciplinary examination (Yes, No)' as dependent variable
Adjusted for: Age

4 Discussion

4.1 Summary of main findings

In this study several differences between the patients assigned to mono- and multidisciplinary examinations were identified.

Patients assigned to multidisciplinary examinations were younger compared to the patients assigned to the monodisciplinary examinations.

The multidisciplinary group were more likely to have beliefs about their pain being related to mental health problems, as well as being more likely to report depression or anxiety. There was also an indication that the patients in the multidisciplinary group were more likely to score above the cutoff of 1.85 on the HSCL-10, which indicates that the patient may suffer from mental health problems; however, this was not statistically significant. The patients in the multidisciplinary group were also more likely to report comorbidities regarding upper back pain, shoulder pain and stomach discomfort.

The patients assigned to multidisciplinary examination were also more likely to use prescription pain medication, while the patients assigned to monodisciplinary examination was more likely to use non-prescription pain medication. Furthermore, the multidisciplinary group received more treatment recommendations, included being recommended follow-up by a physician in primary care and work-related follow-up in primary care.

In addition, the results showed interesting non-differences between the patients assigned to mono- and multidisciplinary examination. The groups were similar reporting on pain indicators, PROMs describing function and almost all work-related variables. A notable result that were not statistically significant is a trend indicating the patients in the multidisciplinary group being more likely to be in the NDI categories moderate and severe/complete disability.

There was also no difference in pain duration between the two groups, as over 90% of the patients in both groups reported pain duration of more than 3 months. This means that both groups consisted of almost only patients who would be classified as patients with chronic pain (10, 18). However, based on recommendations from the clinical guidelines it is expected that most patients referred to specialist health care services are suffering from chronic pain as the guideline recommend referral if pain persist after 6-8 weeks (10).

4.2 Discussion of the main results

An important context in this study is that the groups investigated in this study are two subgroups of all the referred patients referred to the same multidisciplinary specialized clinic. The GPs who refer the patients would most likely consider the two subgroups as the same patient group based on recommendations in the clinical guidelines. Multidisciplinary rehabilitation programs are recommended to patients who have not responded to first-line treatments and who are severely disabled by pain (36), and referral to a multidisciplinary specialized clinics in physical medicine and rehabilitation is recommended if pain persist after 6-8 weeks (10). The author has not been able to find any studies directly comparing patients assigned to mono- and multidisciplinary examinations in a multidisciplinary pain clinic setting.

When assessing the main results of this study there are indications that through the current evaluation process of referrals the senior physicians assign patients with different characteristics to mono- or multidisciplinary examinations at the specialized clinic for physical medicine at UNN. Though there is also uncertainty about the true strength of most associations in this study. There is an overall lack of precision in the results which can be identified through the mostly wide confidence intervals. This means that the true strengths of the associations in this study are unknown. The main reason for the lack of precision is likely the small sample size, as well as for most variables the risk of the event were small (more frequent events would make for a more precise or more narrow confidence interval) (37).

One must also take into consideration when assessing the results that the yellow flags do not operate in isolation and need to be viewed in context (23). It would not be satisfactory to just view the strength of one association in isolation between one of the risk factors and being assigned to mono- or multidisciplinary examinations to describe how well it makes the classification of the patients (38). In order to evaluate the clinical relevance of the results, they need to be assessed in relation to each other; with also variables which are not statistically significant being able to influence conclusions.

Patient characteristics

In this study the patients in the multidisciplinary group were younger than patients in the monodisciplinary group. This difference of almost 6 years in age between the groups can potentially be viewed as a form of discrimination based on age (39); also referred to as "ageism" (40). Neither the national clinical guideline (10) nor the NPGDs (25) mention age as

a factor for treatment recommendations or prioritization of back pain patients. However, there could be other plausible explanations for this observed difference that this study cannot conclusively answer.

One reason for the age difference could be that back pain is less common among a younger population (41). It is well established that back pain affects all age groups (8), from children and adolescents to the elderly. Nevertheless, the prevalence of back pain is at the highest in the age group from 35 to 55 years (41). A systematic review reports that there was around three to four times higher prevalence of chronic LBP at age 50 compared to aged 18-30 (42). It could therefore be perceived as a more serious or urgent problem for a younger person and therefore warranting the allocation of more health resources. In addition, in the general population a strong predictor of having future back pain is having had back pain in the past (8, 43), and knowing that back pain is the largest reason to loss of DALYs (5); assigning more initial resources to younger people could be perceived as more cost-efficient if assumed to prevent future use of health resources due to back pain.

The difference in age could be conscious or unconscious, and could be due to factors on the micro level (such as a clinician's decision-making) or due to factors at the macro level (such as in institutional policies) (44). As mentioned, it is not possible to make any conclusions in this study on the age-difference between the mono- and multidisciplinary group. This is therefore a topic that could be worth examining further in future studies.

Associations between variables describing yellow flags and mono- and multidisciplinary examinations

The findings from this study showed that the differences between the mono- and multidisciplinary group are mainly related to mental health problems. This is in line with research on risk factors or prognostic indicators among back patients as several studies have showed a relationship between the mental health problems among the yellow flags and pain-related outcomes (8, 20, 23). Depressed mood was mentioned as the strongest predictor of chronic neck pain in one study (22). A systematic review from 2015 concluded that symptoms of depression was associated with an increased risk of new episodes of LBP, with the risk being even higher in patients with more severe depression (21). A higher cutoff was chosen for UHI single items in this study as described in 2.8. This separated the patients reporting a higher degree of depression from the patients reporting only some symptoms.

When assessing the ORs and 95% CIs of the variables describing mental health problems (Table 6); the direction of all associations indicate that the multidisciplinary group are more likely to report mental health problems than the monodisciplinary group, while the true strength of the associations are unknown due to the lack of precision in the results. Even though the association between HSCL-10 score >1.85 and mono- and multidisciplinary examinations is not statistically significant, it has the same direction as the other variables describing mental health problems. It is likely that all the mental health variables are clinically relevant when they are viewed in context with each together.

In a referral note from the GP it is quite likely that information of any known mental health problems would be included. The GP could gain knowledge of this by asking directly if the patient feels depressed or experience any other mental health problems like anxiety or stress. In addition, this study shows that the patients in the multidisciplinary group are more likely to report that they believe that their pain is related to mental problems. This strengthens the assumption that the patients may have discussed mental health problems with their GP.

Several comorbidities were also more likely in the multidisciplinary group, as there was a higher likelihood of reporting upper back pain, shoulder pain and stomach discomfort. The prognosis of back pain is worse when it is combined with pain in other parts of the body (8, 10). The clinical relevance of these results (Table 6) are hard to judge, as the precision of the results that are statistically significant are poor and differences were not detected in most of the comorbidities that were investigated. Though it is not likely that a patient would be assigned to a multidisciplinary examination based on only one of these comorbidities being reported in a referral.

There were no differences detected between the groups regarding fear-avoidance beliefs. A systematic review concluded that high fear avoidance beliefs were mainly non-prognostic when used in a population already reporting chronic pain (45). The proportion of patients with chronic pain were over 90% for both groups in this study and could explain the lack of differences observed in fear avoidance beliefs between the mono- and multidisciplinary groups. Other studies have also questioned the usability of the FAB-Q. One study concluded that FAB-Q was not a good measure for fear-avoidance beliefs (46), while another found the FAB-Q to not be a good fit with LBP patients (47). One of the criticisms of the FAB-Q was that the 7-point scoring of the questions were to many, with most questions just supporting

yes or no or a 3-point scale. This could affect the responsiveness of the FAB-Q (46), which means that a difference between the groups could become difficult to detect.

There was a statistically significant higher odds (OR=4.54, CI 95=1.42-13.95) of the patients in the multidisciplinary group having had more than 10 previous sick leaves. The OR is the highest of all the associations made in this study, while at the same time also has the widest CI, which means that there is a great deal of uncertainty of the true strength of this association.

There were no statistically significant associations among the remaining work-related variables and mono- and multidisciplinary examinations in this study, as well as it was difficult to point out any trends. Unfortunately the length of the patients' sick leave was not calculated in this study, as length of sick leave has been suggested as one of the most important work-related yellow flags (10, 48). A date for when the last sick leave started was provided in the dataset, but not the date of the examination, which made a calculation not possible. Length of sick leave is a topic that should be further examined in future studies to shed more light on work-related factors affecting the prognosis or outcomes for the patients.

Associations between variables describing pain and function and mono- and multidisciplinary examinations

There were mostly non-differences between the mono- and multidisciplinary group regarding pain and function. Studies have showed a correlation between reported pain and perceived function levels like the ODI score (49) (50) and the NDI score (51). Other studies though have pointed out that over 40% of the people who report having higher level of chronic pain do not report higher level of disability because of the pain (23).

In this study, neither the ODI or the NDI showed statistically significant differences between the mono- and multidisciplinary groups. However, there was an observable trend where the multidisciplinary group have higher odds of being in the NDI groups with moderate and severe/complete disability (Table 7). This trend is worth noting, as it could have a clinical relevance. Future studies could potentially investigate this further.

In addition, the ODI does not measure any psychological consequences of pain; it is strictly a physical function outcome (29). This means that even though there are differences between the groups regarding mental health problems, there is not necessarily a correlation with ODI scores.

It could also possibly be more difficult for a GP to describe differences in function and pain intensity in this patient group than it would be to list potential mental health problems. All the patients have sought help from the GP because of back pain which was disabling enough to be referred to the hospital, but not everyone would answer yes to feeling depressed or report any other mental health problems. This could conceivably also be a reason why there are a lack of differences between the groups among the PROMs describing function and the pain indicators, while there are differences concerning mental health problems. And as previously mentioned; patients from both groups consist of more than 90% with chronic pain, which could also be a reason why there is a lack of differences between the groups in pain and function.

Patients in the multidisciplinary group were more likely to use prescribed pain medication more than once a week. While not statistically significant, there were also higher odds of patients in the multidisciplinary group reporting daily use of prescribed pain medication. The true strength of these two associations is uncertain due to the lack of precision in the results (Table 7). The GPs are mainly the prescriber of the pain medication to the patients, and therefore it is likely part of the information received in the referral note. When taking all information into account; it is likely prescribed pain medication use have a clinical relevance when assigning patients to either mono- or multidisciplinary examinations.

Interestingly there is a difference in pain medication use even though there were no differences in function and pain intensity. This is interesting because studies have showed an association between pain intensity and strength of pain medication among patients with musculoskeletal pain (52).

Associations between treatment recommendations and mono- and multidisciplinary examinations

Overall, the multidisciplinary group did receive more treatment recommendations than the monodisciplinary group. It is not possible to make any conclusions on reasons for this difference through the available data in this study. However, the fact that the multidisciplinary group received more treatment recommendations strengthens the hypothesis of the multidisciplinary group being more complex. To be able to report more conclusive on this, information from the hospital record would have to be cross-checked or obtained through interviews of both patients and health professionals.

In this study, the patients in the multidisciplinary group are more likely to be recommended follow-up by a physician in primary care, with an OR of 3.1 and a 95% CI of 1.85-5.2. The result is the strongest association made between the groups in this study. The CI is still wide, but even the minimum of 1.85 are likely to have a clinical relevance.

A potential reason for this could be due to the more likely use of prescribed pain medication in the multidisciplinary group. A physician in primary care would be the most likely health professional to follow-up prescribed pain medication use, especially if there were concerns related to this. The message from guidelines for treating back pain is that first choice of therapy should be non-pharmacological (36), though there is a gap between guidelines and real-life practice. The Norwegian clinical guideline for LBP recommends to be restrictive of the use of opioids in the treatment of chronic LBP due to addiction hazard and very little documented effect of opioids as treatment for LBP (10). Studies have also shown that higher doses of prescription opioid doses correlates with a higher use of health care services and worse mental health (53, 54). In addition to administering prescription medication, the GP would also be the one administering the patient's sick leave status, as well as potentially referring the patient to other treatments potentially based on advice from the hospital.

Even though there were no differences in fear-avoidance beliefs, in belief of the pain being due to work, as well as a lack of differences among work-related variables; there were differences in recommendations regarding work-related follow-up. The patients in the multidisciplinary group were more likely to receive a recommendation of work-related follow-up in primary care. However, there is a great deal of uncertainty to the true strength of this result due to lack of precision. It is therefore not possible to make any conclusions regarding work-related follow-up.

From the results of this study it is likely that the patients assigned to multidisciplinary examinations are more complex than patients assigned to monodisciplinary examinations. To explain why there are observed differences between the mono- and multidisciplinary groups, one can assume that a GP would probably provide more information in a referral note concerning a patient he or she perceived as being more complex. This is an unknown which is impossible to account for unless the results from this study were crosschecked with the referral notes.

It is worth noting that research has shown that there is difference in quality of referrals from GPs. There has been some studies in Norway who have investigated the quality of referrals from primary care to medical outpatient clinics (55) or a multidisciplinary outpatient clinic for patients with back pain (56). Both studies concluded that a large share of referrals were lacking in quality or missing relevant information. Another study from Denmark on quality of referrals to hospital departments from the field of nuclear medicine also showed a lack of quality, as well as an increase in quality of the referrals after a feedback response were given to the physicians with low quality referrals (57).

4.3 Methodological discussion

4.3.1 Study design

The cross-sectional design of this study allows to investigate the prevalence of being exposed (in this study the exposures are the different variables in the questionnaires) in the mono- and multidisciplinary groups (which are the outcome in this study). Both exposure and outcome measures are assessed at the same point in time. Therefore, it is not possible to state a causal relationship; if the exposure lead to an outcome or if it happened the other way around (58) (59). Though in this study it is highly unlikely that being assigned to an examination could plausibly lead to an exposure. In this study we are not looking for whether a disease is present or not, but rather the likelihood of being assigned to either a mono- or multidisciplinary examination based on a series of "exposure variables".

The data used in this study is from a national medical quality registry. A registry collecting data on a certain patient group opens several opportunities to investigate characteristics and associations within that patient group, but at the same time there are limitations when using registry data in research (60, 61).

Even though a registry could be designed for use in research, it is not specifically designed to be able to answer a specific research question and aims of a specific study (60). As a result, any non-experimental studies using pre-collected data like registry data to answer a research question could face problems with bias (60). Bias are systematical errors in various parts of the study that could inflate or hide true associations (59) and are commonly divided into three main categories; selection bias, information bias and confounding (60). Nevertheless, if the data in the registry is of good quality and suitable to answer a specific research question, a registry is a good source of information for observational studies.

Using the NNRR is suitable to answer the research question as it covers the whole study population and measures both demographic characteristics and a wide range of clinical characteristics of the study population. If not all relevant outcome measures to answer the aims of the study were collected, it would make it difficult to make conclusions in the study (61). Also, if all relevant covariates have not been collected, it could lead to issues with undetected confounding (61).

This study included only patients from one of the hospitals that the register covers. Including patients from more hospitals would have given a higher number of study subjects that would have increased the statistical power of the study. A larger sample size could potentially increase the precision of the results (37), due to a decrease in the variability of the data. The effect of less variability in the data are narrowing confidence intervals (62). When assessing the results there are mostly wide confidence intervals, which – even if the result is statistically significant – means that there is less certainty about the true effect on the outcome (63).

However, as mentioned in the introduction, the different hospitals differ in the proportion of patients assigned to multidisciplinary examination, ranging from 2,5% in Trondheim to 99,6% in Bergen (28) compared to 11,1% at UNN, indicating that the hospitals have different selection criteria for who should be assigned to a multidisciplinary examination. This means that even though the statistical power of the study would have been increased by including more patients from different hospitals, it would have introduced selection bias not possible to account for.

The findings are therefore only valid in a local context, which in this study is at the specialized clinic for Physical Medicine and rehabilitation at UNN in 2018. By selecting only patients referred only to UNN in this study it becomes difficult to generalize the results of this study, which means the external validity of this study is low (59).

It was also considered to classify the patients in three groups instead of just mono- or multidisciplinary groups. The monodisciplinary group could have been divided into a "examined only by a physician" and "examined only by a physiotherapist". However, due to the low number of patients in the study only two groups were kept. It is reasonable to assume that there also could be differences between patients assigned to an examination by either a physician or a physiotherapist.

4.3.2 Selection bias

Selection bias is a systematic error caused by the study subjects not being representative of the population they came from (60). There is also a distinction between selection bias and the selection of study subjects (59). How the selection of all study subjects is done can affect the external validity of the study, while at the same time not affecting the internal validity of the study. If there are systematic errors in selecting the groups for comparison within the study, the study is subject to selection bias which can affect the study's internal validity. As previously mentioned, excluding patients from the hospitals weakens the external validity of this study, while if included would have had a major negative impact on the internal validity of the study as selection bias would have been introduced.

In previous years the NNRR response rate at UNN was lower; 59% in 2017 and 37% in 2016 (64). If the patients from the previous two years had been included in the study, it would have affected the internal validity of the study. A high number of non-responders would increase the risk of selection bias as non-responders tend to be different in demographic and clinical characteristics from the subjects who decided to be included in a study (60). In this study a good response rate (87% in 2018) increases the internal validity of the study and ensures that the patients included in the study are representative of all the patients examined at the specialized clinic for Physical medicine at UNN in 2018. Nothing is known about the non-responders of the NNRR questionnaires, and therefore it is impossible to know if this group differs from the patients who are in the register.

There is also a possibility that the senior physicians changed practice in how they assigned patients to multidisciplinary examinations over the years, so excluding the use of previous years patients minimize the potential effect of changed practice.

4.3.3 Information bias

An inherent issue with using registry data is information bias, as there will always be questions concerning the quality of the data (61). There could be uncertainty with the accuracy and validity of the data. There could also be uncertainty whether the data is complete or having problems with missing information.

Information bias is a type of systematic error in a study related to how the data is measured (59). When the data is measured erroneously it can lead to misclassification of the individuals in the study (65). Misclassification can be "differential" or "non-differential". When a

misclassification is more likely to occur in one group than in the group which it is compared to, it leads to differential misclassification (60). This could lead to an assumed association between the exposure and outcome even if there wasn't a true association. When misclassification is non-differential the misclassification is not more likely in any of the groups, and potentially as a result of a problem in how the data is collected. Non-differential misclassification normally leads to odds ratios being shifted towards 1, which means that a true association potentially is not detected.

Recall bias

A type of information bias which often is a problem with self-reporting questionnaires like the first patient questionnaire, is recall bias. Recall bias occurs when the study participants are not able to recall the information they are asked for (59). Information could be over- or underreported in one or both groups being compared.

Having a disease could lead to overreporting exposures due to being more likely to try to find an explanation to the disease (59). This would lead to an overestimation of the association between the exposure and outcome. In this study the patients in multidisciplinary group and the patients in the monodisciplinary group are overall the same disease category, as all patients are back pain patients. It is unlikely that patients would answer differently in the first patient questionnaire based on which type of examination they were assigned to. The first patient questionnaire is filled out prior to the examination, and a patient assigned to a multidisciplinary examination does not receive any other information from the hospital than what a patient assigned to a monodisciplinary examination does. This means that the risk of differential recall is low when considering the data sourced from the first patient questionnaire.

Underreporting could be due to not being able to recall the information the person is asked to report. Most commonly underreporting leads to non-differential misclassification and a true association could be missed as a result. Naturally it is more difficult to remember something that happened a long time ago. This type of recall bias can be reduced through how the questionnaire is designed. One option is to limit the time period from which the person is asked about (65). Several of the PROMs used in the NNRR first patient questionnaire uses this method of limiting the recall period. The ODI, the NDI and EQ-5D-3L all ask how the current function of the patient is, while the HSCL-10 have a recall period of the last seven days and the UHI of the last four weeks.

The question asking for number of previous sick leaves has several validity problems; this variable is sourced from the health professional questionnaire, which means the health professional is required to have asked the patient directly about number of sick leaves or obtained this information through the referral note or the hospital record. It is therefore subjected to recall bias if sourced from the patient, as he or she would have to remember over several years to estimate the number of sick leave due to back pain. Due to all these factors in combination with the previously discussed lack of precision in the OR, one must be cautious of making any conclusions based on this result.

Social desirability bias

Social desirability bias can also lead to underreporting, as well as overreporting (65). This type of bias is more likely in variables with private or sensitive information. In this study questions about pain medication use could be affected by this bias, and potentially underreported as a result. It is not likely that this would be more underreported by either the mono- or multidisciplinary group. However, if overall underreporting is present for pain medication use, it will lead to non-differential misclassification. This potentially underestimates the strength of the association of pain medication use and being assigned to the mono- or multidisciplinary group.

Confirmation bias

Confirmation bias could affect some of the information sourced from the health professional questionnaire. Confirmation bias is a psychological bias where a person makes a decision based on preconceived beliefs (65). The patients assigned to multidisciplinary examination have at least two health professionals assessing their case. Since the patient already have been selected to a more resource-demanding examination on the assumption of being “more complex”, he/she could be more likely to receive a treatment recommendation just because of that preconceived assumption. This would mean that there is a risk of differential misclassification of answers from health professional questionnaire regarding treatment recommendations. However, when considering the other results of this study, confirmation bias is not very likely to be the main reason for differences in treatment recommendations, as there are other more plausible reasons for more treatment recommendations in the multidisciplinary group.

Item nonresponse bias

Missing data could lead to "Item nonresponse bias". Item nonresponse is the cases where a study subject has left out answering questions for various reasons, which could also lead to misclassification. Item nonresponse does not automatically cause bias. Substantial item nonresponse bias is related to the quantity of item nonresponse and correlation with the outcome measure. This means potential bias is only related to the specific outcome measure, and not the entire study (66). When assessing item nonresponse, it is important to evaluate if it is due to the exposure or the outcome to decide if the misclassification is differential or non-differential (60). In this study it is not likely missing information is due to the outcome, due to the same reasoning as for potential differential recall. In this study item nonresponse have been counted for every exposure variable and listed separately in the descriptive statistics, as well as investigated manually and given reason for in the results if possible. In the association analysis of this study using binary logistic regression, patients with missing information in any of the exposure variables in the model are excluded (60).

In this study there were notable high numbers of nonresponse in the questions regarding non-prescription and prescription pain medication use. After review, it was apparent that almost all patients who didn't provide information on either non-prescription or prescription pain medication use, did provide information on the other. So, this is missing information related to the exposure, and therefore leading to a non-differential misclassification which in this case shift the association towards 1 (sensitivity analysis using imputation of new variables with assumed answers to the questions confirmed this).

Another notable high number of nonresponses were in the question of "not felt wanted back by your employer". There were also high numbers of nonresponses in the other work-related questions, mostly due to non-responders not having current employment, though substantively lower than for "not felt wanted back by employer". This could be a case of social desirability bias leading to nonresponse. It is not unlikely that the non-responder feels that the question is sensitive and do not want to disclose displeasure with the employer.

Validity of outcome measures

Also affecting the quality of the information from the questionnaire is the validity and reliability of the different PROMs and other variables.

In a systematic review from 2011 ODI was identified as the most used outcome measure of function in chronic LBP-patients and was found to be valid and reliable (67). A review of the culturally adapted versions of the ODI stated that researchers can be confident in the validity and internal consistency of the test (68). Though another systematic review form 2018 suggest that there could be inadequate evidence for content validity of the ODI, and raise doubt about the use of the total score of ODI (69).

A systematic review from 2018 evaluating PROMs used for neck disorders found the NDI to have moderate to excellent reliability and moderate to good validity (70).

Pain VAS was found to be reliable in a chronic LBP population and referred to as the "gold standard" for measuring pain in a chronic LBP population (67).

FAB-Q was stated as the most used psychosocial outcome measure and found to be reliable and valid in a chronic LBP population in one systematic review (67). However, later studies have raised concerns about the validity of the FAB-Q. A systematic review concluded that high fear avoidance beliefs were mainly non-prognostic when used in a population already reporting chronic pain (45), which over 90% of the patients in both groups in this study suffer from. Another study from 2020 concluded that FAB-Q was not a good measure of fear-avoidance beliefs and that the 7 categories for each question were too many, and possibly affecting the responsiveness of the FAB-Q (46). Another study found that FAB-Q was to not a good fit with LBP patients and also state that the two sub scores should not be summarized (47).

There are no systematic reviews evaluating the validity of the UHI, but an cross-sectional study comparing LBP-patients with the general population concluded that the LBP-patients reported more subjective health complaints than the general population (71).

HSCL-10 was checked for correlation with several other commonly used tools for evaluating mental health, and was found to strongly correlate with the more comprehensive HSCL-25 and validated through high associations with other validated tools (72).

Patients with LBP have been found out to score the EQ-5D-3L differently compared to the general population (73). A systematic review suggested using caution when assuming the validity of the EQ-5D-3L and EQ-VAS in LBP patients (74).

4.3.4 Interrater reliability

Another uncertainty in this study is interrater reliability between the health professionals. Interrater reliability is to what degree different observers make the same judgements of the same phenomenon (75). In this study it would be to what degree the senior physicians or the health professionals come to the same conclusions when assessing similar patients.

No studies have been done regarding interrater reliability between the senior physicians who assign patients to the examinations. For every single patient different levels of information could be available when the referral of the patient is evaluated (Figure 1). If the senior physicians evaluate referrals differently it could lead to selection bias. An example could be that one senior physician had a substantial lower threshold for assigning younger patients to multidisciplinary examinations than the other senior physicians. Data from referrals and hospital records are needed to account for this, as well as potentially from the senior physicians as well.

There is a written guideline describing how to fill out the health professional questionnaire available to the health professionals performing the examinations, but no studies have been done to assess the interrater reliability between the health professionals in this context. Variables sourced from the health professional questionnaire regarding work and sick leave status, diagnosis and treatment recommendations must be interpreted with this uncertainty in mind, and they could be subject to information bias as a result. Work and sick leave status could have been self-reported through the first patient questionnaire, which would have eliminated interrater reliability issues for those variables.

4.3.5 Confounding

Confounding occurs when a factor is both associated with the exposure and the outcome, but are not in the causal pathway between them (59). This means that if a confounder is present, the observed association between the exposure and outcome can be different from the true association between the exposure and the outcome. Confounders can be accounted for through statistical analysis. However, residual confounding would remain (76).

In this study age was identified as a confounder. When comparing the patient characteristics of the mono- and multidisciplinary group it was clear that the group assigned to multidisciplinary examination had a statistically significant lower mean age. Age would not be in the causal pathway between exposure variables and being assigned to a

multidisciplinary examination or not (outcome). It is reasonable though to assume that the senior physicians would have a lower threshold to assign more resources to a younger patient. This means that it was important to find out if age confounded the associations between the different variables and mono- or multidisciplinary examination and adjust for this potential confounding effect to be able to determine a true association. Sensitivity analysis showed that age also was associated with several of the exposure variables.

To reduce the effect of confounding due to age, the association analysis was performed using a binary logistic regression model adjusting for age.

Sex were also considered as a confounder as sex could be associated with several of the exposure variables. The proportions of men and women in the groups were almost identical in the mono- and multidisciplinary group, and almost identical with the proportion of men and women in the overall study sample. So, there was no indication that sex was associated with the outcome and therefore not affecting the decision-making in assigning patients to either mono- or multidisciplinary examinations, which meant that sex would not confound the results.

There was a notable difference between the groups in smoking status, with a higher proportion of smokers in the monodisciplinary group. Smoking status are likely associated with several of the exposure variables. However, it is not likely that smoking status are associated with the outcome. Smoking status should not affect the decision-making of being assigned to mono- or multidisciplinary examinations. There were likely other reasons for the higher proportion of smokers in the monodisciplinary group; with higher age being the most likely reason.

4.3.6 Strengths and limitations

The main strength of this study is the high participation rate which increases the internal validity of the study (59, 60). There is a low risk of selection bias as the sample size consist of almost all patients examined at UNN eligible to participate in the NNRR. Another strength of the study is the broad range of information from both patients and health professionals involved sourced through the NNRR questionnaires, which reduces the risk of relevant outcome measures or confounders being missed.

The main limitation of the study is the small sample size which contribute to the lack of precision that can be seen in the results, leading to uncertainty about the true strength of

associations between the variables and the mono- and multidisciplinary group (37). In addition, just using data from a single location is limits the external validity of the results, making it very difficult to generalize about the process of assigning patients to mono- or multidisciplinary examinations (59).

Another limitation is in the classification of the patients in this study. The monodisciplinary group could have been divided into separate groups according to health profession present at the first consultation. It is plausible that patients assigned to physicians could differ from patients assigned to physiotherapists.

There are several inherent limitations due to the nature of the study design and with data being predominantly self-reported (59, 60, 65). The cross-sectional design makes it not possible to say something definitive about the causal relationships between exposure and outcome; though it is unlikely that being assigned to a multidisciplinary examination would lead to an exposure variable. The data have some problems regarding information bias as discussed in 4.3.3; with underreporting due to various reasons being the main concern. There could be underreporting due to recall bias, social desirability bias and item nonresponse bias. All are likely leading to non-differential misclassification in different variables in this study, which weakens the true strength of associations between the variables and the mono- and multidisciplinary groups. A strength of this study regarding information bias is the little risk of differential misclassification, as it is not likely the mono- and multidisciplinary groups would answer differently in the first patient questionnaire.

Though there are limitations of using self-reported data, a strength of the NNRR is that it uses commonly used PROMs and outcome measures regarding pain discussed in 4.3.3, which are mostly validated in a back pain population. This means that results from studies using information from the NNRR are easily compared to most available research in the field. However, recent literature has stated that the FAB-Q most likely is insufficient in measuring fear-avoidance beliefs (45-47). This means that the NNRR could have a void regarding usable data concerning fear-avoidance beliefs. Recent literature has also advised using caution when assuming the validity of the EQ-5D-3L and EQ-VAS in an LBP population (74), as LBP patients seem to score the EQ-5D-3L differently than the general population (73).

Another limitation lies in the information which is not known. To gain further knowledge of the patients assigned to mono- or multidisciplinary examinations one would need information

provided in the referrals, information on the evaluation of referrals from senior physicians and information from hospital records, as well as information from subsequent questionnaires from the NNRR to see if there were differences in long term outcomes between the groups. Without this additional information it is not possible to conclude if the patients assigned to multidisciplinary examinations in this study are indeed the "right patients".

Also, there is a limitation regarding the health professional questionnaire. It does not provide any information of which risk factors the health professional(s) believe to affect the patient's back pain. A short multiple choice in the health professional questionnaire with risk factors (like mental health problems, work-related problems, etc.) could potentially have provided more insight into a potential multifactorial condition like back pain. If this information was available, it could have strengthened the indication in this study of the multidisciplinary group being more likely to report mental health problems.

This study only brings a snapshot of a practice at UNN. Doing the same calculations in successive years would make it possible to see trends and compare it with changing clinical practices or decision-making by the senior physicians. In addition, by summarizing NNRR data from successive years after 2018 one would gain statistical power enabling a higher degree of precision in the results and more certainty of the true strength of the associations between the different variables and being assigned to mono- or multidisciplinary examinations.

4.4 Implications for practice

This study brings new information about the evaluation of referrals at the specialized clinic for Physical medicine at UNN and is the first study in Norway investigating differences between patients assigned to mono- or multidisciplinary examinations for back pain conditions. This study has shed light and filled in new knowledge which have not been investigated before.

One of the stated purposes of the NNRR is to aid the process of assigning the right patients to the right interventions. As mentioned, this study cannot determine if "the right patients" are assigned to multidisciplinary examinations. However, by using the information from this study, the senior physicians at UNN can evaluate their protocol for assigning patients to multidisciplinary examinations and evaluate if their process works as intended. Also, this

study could give the senior physicians an indication of what type of information they should require in addition from GP's referrals.

Another stated purpose of the NNRR is to improve quality locally and eventually nationally. By using the information from this study, UNN can potentially improve upon their interventions. The indication of patients assigned to multidisciplinary examinations suffering from more mental health problems could validate extended use of health professionals with a specialty in mental health in this setting. This could either be psychologist or a physiotherapist or physician trained in cognitive behavior treatment. Studies have showed that identifying risk factors for mental health problems is an important step to psychological interventions, however it would be ineffective (not cost-effective) if these interventions were applied to all patient regardless of risk status (23).

This study could also be used in a future cohort study investigating differences between the examinations themselves. Prospective data from either the NNRR or another registry like the sick leave registry of the Norwegian Labor and Welfare Administration could be used to evaluate the if patients assigned to mono- or multidisciplinary examinations had different long term outcomes regarding pain, physical and mental disability or return to work. Such a study would gain knowledge of the effectiveness and usefulness of multidisciplinary examinations.

Information from this study could also be used to improve the quality of the NNRR. In the health professional questionnaire, there is no possibility to register if a psychologist or other health professional with formal training in mental health care was present at the first consultation. We would recommend more options for registering which health professional was present, as it would make it possible to evaluate the multidisciplinary team or targeted interventions more extensively.

It would also be beneficial if the NNRR provided the number of days of sick leave, as length of sick leave is a predictor for patient outcome. The health professional questionnaire currently provides demographic information like employment status, sick leave status and disability pension status. However, if the patient self-reported this information in the first patient questionnaire, the information would potentially be of better quality as it comes directly from the patient. In addition, for research purposes more information could be

collected directly from the patient, and reduce the need having access to information from the health professional questionnaire.

5 Conclusion

This study shows that patients assigned for multidisciplinary examinations at UNN are younger than patients assigned for monodisciplinary examinations. There are also indications that patients assigned to multidisciplinary examinations suffers from more mental health problems, as well as being more likely to use pain medication requiring a prescription. The patients assigned to multidisciplinary examinations were also receiving overall more treatment recommendations from the health professionals and were especially more likely to be recommended to follow-up by a physician in primary care. There were no differences between the groups concerning pain, function and work-related variables.

Figures and Tables

References

1. Hoy D, Bain C, Williams G, March L, Brooks P, Blyth F, et al. A systematic review of the global prevalence of low back pain. Arthritis & Rheumatism. 2012;64(6):2028-37.
2. Ihlebæk C, Brage S, Natvig B, Bruusgaard D. Forekomst av muskelog skjelettlidelser i Norge. Tidsskrift for Den norske legeforening. 2010;130(23):2365.
3. Hogg-Johnson S, van der Velde G, Carroll LJ, Holm LW, Cassidy JD, Guzman J, et al. The burden and determinants of neck pain in the general population: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. Spine (Phila Pa 1976). 2008;33(4 Suppl):S39-51.
4. Nygaard ØP, Rø M, Inge Andersen T, Zwart J-A. Nakkesmerter med og uten nerverotsaffeksjon. Tidsskrift for den Norske legeforening. 2010;130(22):2252.
5. Knudsen AK, Tollånes MC, Haaland ØA, Kinge JM, Skirbekk V, Vollset SE. Sykdomsbyrde i Norge 2015. Resultater fra Global Burden of Diseases, Injuries, and Risk Factors Study 2015 (GBD 2015). 2017.
6. Kyu HH, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2018;392(10159):1859-922.
7. Lærum E, Brage S, Ihlebæk C, Johnsen K, Natvig B, Aas E. Et muskel- og skjelettregnskap. 2013.
8. Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, et al. What low back pain is and why we need to pay attention. The Lancet. 2018;391(10137):2356-67.
9. Deyo RA, Weinstein JN. Low back pain. The New England Journal of Medicine. 2001.
10. Lærum E, Brox JI, Storheim K, Espeland A, Haldorsen E, Munch-Ellingsen J, et al. Nasjonale kliniske retningslinjer for korsryggsmerter - med og uten nerverotsaffeksjon. 2007.
11. Guzman J, Haldeman S, Carroll LJ, Carragee EJ, Hurwitz EL, Peloso P, et al. Clinical Practice Implications of the Bone and Joint Decade 2000–2010 Task Force on Neck Pain and Its Associated Disorders: From Concepts and Findings to Recommendations. Spine. 2008;33(4S):S199-S213.
12. Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. The Lancet. 2018;391(10137):2368-83.
13. Scascighini L, Toma V, Dober-Spielmann S, Sprott H. Multidisciplinary treatment for chronic pain: a systematic review of interventions and outcomes. Rheumatology. 2008;47(5):670-8.
14. van Middelkoop M, Rubinstein SM, Kuijpers T, Verhagen AP, Ostelo R, Koes BW, et al. A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. European Spine Journal. 2011;20(1):19-39.
15. Kamper SJ, Apeldoorn AT, Chiarotto A, Smeets RJ, Ostelo R, Guzman J, et al. Multidisciplinary biopsychosocial rehabilitation for chronic low back pain. Cochrane Database of Systematic Reviews. 2014(9).
16. Oliveira CB, Maher CG, Pinto RZ, Traeger AC, Lin C-WC, Chenot J-F, et al. Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. European Spine Journal. 2018;27(11):2791-803.

17. Burton AK, Balagué F, Cardon G, Eriksen H, Henrotin Y, Lahad A, et al. European guidelines for prevention in low back pain: November 2004. European Spine Journal. 2006;15(Suppl 2):s136.
18. Rozenberg S. Chronic low back pain: definition and treatment. Rev Prat. 2008;58(3):265-72.
19. Pincus T, Burton AK, Vogel S, Field AP. A Systematic Review of Psychological Factors as Predictors of Chronicity/Disability in Prospective Cohorts of Low Back Pain. Spine. 2002;27(5):E109-E20.
20. Pincus T, McCracken LM. Psychological factors and treatment opportunities in low back pain. Best Practice & Research Clinical Rheumatology. 2013;27(5):625-35.
21. Pinheiro MB, Ferreira ML, Refshauge K, Ordoñana JR, Machado GC, Prado LR, et al. Symptoms of Depression and Risk of New Episodes of Low Back Pain: A Systematic Review and Meta-Analysis. Arthritis Care & Research. 2015;67(11):1591-603.
22. Shahidi B, Curran-Everett D, Maluf KS. Psychosocial, Physical, and Neurophysiological Risk Factors for Chronic Neck Pain: A Prospective Inception Cohort Study. The Journal of Pain. 2015;16(12):1288-99.
23. Nicholas MK, Linton SJ, Watson PJ, Main CJ, Group tDotFW. Early Identification and Management of Psychological Risk Factors ("Yellow Flags") in Patients With Low Back Pain: A Reappraisal. Physical Therapy. 2011;91(5):737-53.
24. Aktuell informasjon om lov og forskrift for prioriteringsveilederne: Helsedirektoratet; [updated 2019 September 20.; cited 2020 September 18.]. Available from: https://www.helsedirektoratet.no/veiledere/prioriteringsveiledere/aktuell-informasjon-om-lov-og-forskrift-for-prioriteringsveilederne.
25. Fysikalsk medisin og rehabilitering - Prioriteringsveileder: Helsedirektoratet; [updated 2018 August 16.; cited 2020 September 10.]. Available from: https://www.helsedirektoratet.no/veiledere/prioriteringsveiledere/fysikalsk-medisin-og-rehabilitering.
26. Norsk nakke- og ryggregister: Nasjonalt servicemiljø for medisinske kvalitetsregistre; [cited 2020 September 9.]. Available from: https://www.kvalitetsregistre.no/registers/norsk-nakke-og-ryggregister.
27. Om medisinske kvalitetsregistre: Nasjonalt servicemiljø for medisinske kvalitetsregistre; [cited 2020 September 18.]. Available from: https://www.kvalitetsregistre.no/om-medisinske-kvalitetsregistre-0.
28. Anke A, Granan LP. Norsk nakke- og ryggregister (NNRR): Årsrapport for 2018 med plan for forbedringstiltak. Universitetssykehuset Nord-Norge, Oslo Universitetssykehus; 2019 30. September 2019.
29. Fairbank JC, Pynsent PB. The Oswestry disability index. Spine. 2000;25(22):2940-53.
30. Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. Journal of manipulative and physiological therapeutics. 1991.
31. Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins Symptom Checklist (HSCL): A self‐report symptom inventory. Behavioral science. 1974;19(1):1-15.
32. Kleppang AL, Hagquist C. The psychometric properties of the Hopkins Symptom Checklist-10: a Rasch analysis based on adolescent data from Norway. Family Practice. 2016;33(6):740-5.
33. Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain. 1993;52(2):157-68.
34. Eriksen H, Ihlebæk C, Ursin H. A scoring system for subjective health complaints (SHC). Scandinavian journal of public health. 1999;27:63-72.
35. Foundation ER. EQ-5D-3L User Guide. EuroQol Research Foundation; 2018.

36. Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. The Lancet. 2018;391(10137):2368-83.
37. Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Chapter 15: Interpreting results and drawing conclusions. Cochrane handbook for systematic reviews of interventions: John Wiley & Sons; 2019. p. 403-31.
38. Mallen CD, Peat G, Thomas E, Dunn KM, Croft PR. Prognostic factors for musculoskeletal pain in primary care: a systematic review. British Journal of General Practice. 2007;57(541):655-61.
39. Krieger N. Discrimination and Health Inequities. International Journal of Health Services. 2014;44(4):643-710.
40. Developing an ethical framework for healthy ageing: report of a WHO meeting, Tübingen, Germany, 18 March 2017. Geneva: World Health Organization; 2017.
41. Freburger JK, Holmes GM, Agans RP, Jackman AM, Darter JD, Wallace AS, et al. The rising prevalence of chronic low back pain. Arch Intern Med. 2009;169(3):251-8.
42. Meucci RD, Fassa AG, Faria NMX. Prevalence of chronic low back pain: systematic review. Rev Saude Publica. 2015;49:1-.
43. Dunn KM, Hestbaek L, Cassidy JD. Low back pain across the life course. Best Practice & Research Clinical Rheumatology. 2013;27(5):591-600.
44. Wyman MF, Shiovitz-Ezra S, Bengel J. Ageism in the Health Care System: Providers, Patients, and Systems. In: Ayalon L, Tesch-Römer C, editors. Contemporary Perspectives on Ageism. Cham: Springer International Publishing; 2018. p. 193-212.
45. Wertli MM, Rasmussen-Barr E, Weiser S, Bachmann LM, Brunner F. The role of fear avoidance beliefs as a prognostic factor for outcome in patients with nonspecific low back pain: a systematic review. The Spine Journal. 2014;14(5):816-36.e4.
46. Aasdahl L, Marchand GH, Gismervik SØ, Myhre K, Fimland MS, Røe C. The Fear Avoidance Beliefs Questionnaire (FABQ) Does it Really Measure Fear Beliefs? Spine. 2020;45(2):134-40.
47. Meroni R, Piscitelli D, Bonetti F, Zambaldi M, Guccione AA, Pillastrini P. Rasch analysis of the Italian version of fear avoidance beliefs questionnaire (FABQ-I). Disability and Rehabilitation. 2015;37(2):151-7.
48. Waddell G, Burton AK. Occupational health guidelines for the management of low back pain at work: evidence review. Occupational Medicine. 2001;51(2):124-35.
49. MacCormick AP, Sharma H. Does the severity of pain correlate with severity of functional disability? Factors influencing 'patient reported outcome measures' in spinal patients. SICOT J. 2018;4:43-.
50. Hong S, Shin D. Relationship between pain intensity, disability, exercise time and computer usage time and depression in office workers with non-specific chronic low back pain. Medical Hypotheses. 2020;137:109562.
51. Hains F, Waalen J, Mior S. Psychometric properties of the neck disability index. J Manipulative Physiol Ther. 1998;21(2):75-80.
52. Muller S, Bedson J, Mallen CD. The association between pain intensity and the prescription of analgesics and non-steroidal anti-inflammatory drugs. Eur J Pain. 2012;16(7):1014-20.
53. Kobus AM, Smith DH, Morasco BJ, Johnson ES, Yang X, Petrik AF, et al. Correlates of higher-dose opioid medication use for low back pain in primary care. J Pain. 2012;13(11):1131-8.
54. Morasco BJ, Yarborough BJ, Smith NX, Dobscha SK, Deyo RA, Perrin NA, et al. Higher Prescription Opioid Dose is Associated With Worse Patient-Reported Pain Outcomes and More Health Care Utilization. J Pain. 2017;18(4):437-45.

55. Lønning KJ, Kongshavn T, Husebye E. Quality of referrals from general practitioners to medical polyclinics. Tidsskr Nor Laegeforen. 2009;129(18):1868-9.
56. Gulati S, Jakola AS, Solheim O, Gabrielsen S, Godø PA, Skogstad B, et al. Assessment of referrals to a multidisciplinary outpatient clinic for patients with back pain. J Man Manip Ther. 2012;20(1):23-7.
57. Grupe P, Grinsted P, Møldrup M, Høilund-Carlsen PF. Quality improvement of referrals from practising physicians. An example of a practicable strategy. Ugeskr Laeger. 2006;168(14):1434-8.
58. Setia MS. Methodology Series Module 3: Cross-sectional Studies. Indian J Dermatol. 2016;61(3):261-4.
59. Gordis L. Epidemiology. 5th ed. Philadelphia, PA: Elsevier Saunders; 2014.
60. Psoter KJ, Rosenfeld M. Opportunities and pitfalls of registry data for clinical research. Paediatric Respiratory Reviews. 2013;14(3):141-5.
61. Gliklich RE, Leavy MB, Dreyer NA. Patient registries. Registries for Evaluating Patient Outcomes: A User's Guide [Internet] 4th edition: Agency for Healthcare Research and Quality (US); 2020.
62. Biau DJ, Kernéis S, Porcher R. Statistics in brief: the importance of sample size in the planning and interpretation of medical research. Clin Orthop Relat Res. 2008;466(9):2282-8.
63. Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. British Medical Journal (Clinical research ed). 1986;292(6522):746-50.
64. Norsk nakke- og ryggregister: Resultater: Nasjonalt servicemiljø for medisinske kvalitetsregistre; [cited 2020 November 11.]. Available from: https://www.kvalitetsregistre.no/registers/532/resultater.
65. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. J Multidiscip Healthc. 2016;9:211-7.
66. Elliott MN, Edwards C, Angeles J, Hambarsoomians K, Hays RD. Patterns of unit and item nonresponse in the CAHPS Hospital Survey. Health Serv Res. 2005;40(6 Pt 2):2096-119.
67. Chapman JR, Norvell DC, Hermsmeyer JT, Bransford RJ, DeVine J, McGirt MJ, et al. Evaluating Common Outcomes for Measuring Treatment Success for Chronic Low Back Pain. Spine. 2011;36:S54-S68.
68. Sheahan PJ, Nelson-Wong EJ, Fischer SL. A review of culturally adapted versions of the Oswestry Disability Index: the adaptation process, construct validity, test-retest reliability and internal consistency. Disabil Rehabil. 2015;37(25):2367-74.
69. Chiarotto A, Ostelo RW, Boers M, Terwee CB. A systematic review highlights the need to investigate the content validity of patient-reported outcome measures for physical functioning in patients with low back pain. Journal of Clinical Epidemiology. 2018;95:73-93.
70. Bobos P, MacDermid JC, Walton DM, Gross A, Santaguida PL. Patient-Reported Outcome Measures Used for Neck Disorders: An Overview of Systematic Reviews. Journal of Orthopaedic & Sports Physical Therapy. 2018;48(10):775-88.
71. Hagen EM, Svensen E, Eriksen HR, Ihlebæk CM, Ursin H. Comorbid Subjective Health Complaints in Low Back Pain. Spine. 2006;31(13):1491-5.
72. Schmalbach B, Zenger M, Tibubos AN, Kliem S, Petrowski K, Brähler E. Psychometric Properties of Two Brief Versions of the Hopkins Symptom Checklist: HSCL-5 and HSCL-10. Assessment. 2019;0(0):1073191119860910.
73. van Dongen JM, van denBerg B, Bekkering GE, van Tulder MW, Ostelo RWJG. Patient versus general population health state valuations: a case study of non-specific low back pain. Qual Life Res. 2017;26(6):1627-33.

74. Chiarotto A, Terwee CB, Kamper SJ, Boers M, Ostelo RW. Evidence on the measurement properties of health-related quality of life instruments is largely missing in patients with low back pain: A systematic review. Journal of Clinical Epidemiology. 2018;102:23-37.
75. Fray B. The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation. 2018.
76. Flanders WD, Klein M, Darrow LA, Strickland MJ, Sarnat SE, Sarnat JA, et al. A method for detection of residual confounding in time-series and other observational studies. Epidemiology. 2011;22(1):59-67.

Appendix

ICD-10 diagnosis code overview

Category	ICD10 Code	Term
Non-specific LBP	M545	Lumbago
Non-specific neck pain	M542	Smerte i nakke
Other or multiple diagnosis	M059	Uspesifisert seropositiv reumatoid artritt
	M060	Seronegativ reumatoid artritt
	M169	Uspesifisert hofteleddsartrose
	M255	Leddsmerte
	M2561	Stivhet i ledd, ikke klassifisert annet sted (?)
	M418	Andre spesifiserte former for skoliose
	M419	Uspesifisert skoliose
	M431	Spondylolistese
	M4312	Spondylolistese (?)
	M45	Ankyloserende spondylitt
	M461	Sakroiliitt, ikke klassifisert annet sted
	M463	Infeksjon i mellomvirvelskive
	M471	Annen spondylolyse med myelopati
	M472	Annen spondylose med radikulopati
	M478	Annen spesifisert spondylose
	M4782	Annen spesifisert spondylose (?)
	M4792	Uspesifisert spondylose (?)
	M480	Spinal stenose
	M4806	Spinal stenose (?)
	M501	Lidelse i cervikalskive, med radikulopati
	M511	Lidelser i lumbalskive og andre mellomvirvelskiver, med radikulopati
	M513	Annen spesifisert degenerativ lidelse i mellomvirvelskive
	M5385	Andre spesifiserte rygglidelser
	M541	Radikulopati
	M543	Isjialgi
	M544	Lumbago med isjialgi
	M546	Smerte i torakaldelene
	M548	Annen spesifisert ryggsmerte
	M549	Uspesifisert ryggsmerte
	M706	Trokanterbursitt
	M751	«Rotator cuff syndrome»
	M755	Bursitt i skulder
	M759	Uspesifisert skulderlidelse
	M760	Gluteal tendinitt
	M771	Lateral epikondylitt
	M790	Uspesifisert reumatisme
	M791	Myalgi
	M7911	Myalgi (?)
	M7915	Myalgi (?)
	M7919	Myalgi (?)
	M796	Smerte i ekstremitet
	M7961	Smerte i ekstremitet (?)
	M7965	Smerte i ekstremitet (?)
	M797	Fibromyalgi
	M940	Smerte i overgang mellom ribben og ribbensbrusk
	S122	Brudd i annen spesifisert cervikalvirvel
	S220	Brudd i torakalvirvel
	S320	Brudd i lumbalvirvel
	D391	Eggstokk (svulst)
	G439	Uspesifisert migrene
	G442	Tensjonshodepine
	G809	Uspesifisert cerebral parese
	R42	Svimmelhet
	R51	Hodepine
	R073	Annen brystsmerte
	Q760	Spina bifida occulta
	T913	Følgetilstander etter skade på ryggmarg
	E669	Uspesifisert fedme
	M545 + M542	“Lumbago” in combination with “Smerte i nakke”

	M545 + any M542 + any	"Lumbago" in combination with any other diagnosis "Smerte i nakke" in combination with any other diagnosis

Radiological findings

There were no statistically significant differences in the listed radiological findings between the patients in the multidisciplinary group and the monodisciplinary group.

Radiological findings

	Examination type		p-value(2)
	Mono-disciplinary	Multi-disciplinary	
MRI Findings			
Undergone MRI, N (%)	487 (83.7)	58 (79.5)	0.363(1)
Normal, N (%)	187 (32.1)	28 (38.4)	0.286(1)
Disc herniation, N (%)	162 (27.8)	23 (31.5)	0.511(1)
Spinal stenosis, N (%)	45 (7.7)	2 (2.7)	0.085(1)
Recess stenosis, N (%)	88 (15.1)	13 (17.8)	0.549(1)
Modic changes, N (%)	70 (12.0)	6 (8.2)	0.228(1)

1 Chi-Square Test, 2 Fischer's exact p-value listed for variables with groups containing less than 10 patients