Statistical analysis plan (SAP) for:

Modifiable prognostic factors of high costs related to healthcare utilization and productivity loss among people on sick leave due to musculoskeletal disorders - an external validation study (working title)

Project	The MI-NAV project
NCT number	NCT03871712
Document date	21.08.2023

Introduction

Scope

This document is a supplement to the MI-NAV protocol (ClinicalTrials.gov Identifier: NCT03871712) and comprises a statistical analysis plan (SAP) for the article "Modifiable prognostic factors of high costs related to healthcare utilization and productivity loss among people on sick leave due to musculoskeletal disorders - an external validation study". The current SAP has been written after data collection was finished. However, the SAP will be uploaded to ClinicalTrials.gov before we enter the study database for the subsequent analyses.

Administrative information

Version of SAP

1.0

Study sponsor

This study is part of the MI-NAV project, a large-scale project funded by the Research Council of Norway, through the program "Sickness absence, work, and health" (280431/GE).

Contributors to SAP

Name	Title	Role
Margreth Grotle[1,2]	Professor	Principal Investigator
Rikke Munk Killingmo[1]	PhD	Main author of SAP
Britt Elin Øiestad[1]	Professor	Contributor to SAP
Anne Therese Tveter[3,1]	Ass. Professor	Contributor to SAP
Esther Maas[4]	Ass. Professor	Contributor to SAP
Are Hugo Pripp[2,1]	Professor	Statistical advisor
Fiona Aanesen[5]	PhD	Contributor to SAP
Alexander Tingulstad[1]	PhD-student	Contributor to SAP
Tarjei Rysstad[1]	PhD-student	Contributor to SAP

[1]Oslo Metropolitan University, [2]Oslo University Hospital, [3]Diakonhjemmet Hospital, [4]Vrije University Amsterdam [5]STAMI National Institute of Occupational Health

Signatures

I hereby declare that I have reviewed and approved the statistical analysis plan for this study.

Rikke Munk Killingmo
Main author of the SAP

Margreth Grotle
Principal Investigator

Study aim

The primary aim of this study is to assess the external validity of identified associations found between predefined modifiable prognostic factors and high costs related to healthcare utilization and productivity loss among people on sick leave due to musculoskeletal disorders, in the work package 3 (WP3) of the MI-NAV project. A secondary aim is to assess the external validity of identified associations found between predefined modifiable prognostic factors and high costs related to separately 1) healthcare utilization, and 2) productivity loss.

Method

Design and setting

This study contains secondary analyses embedded in a three-arm, pragmatic randomised controlled trial (RCT) with 6 months of follow-up, conducted within the Norwegian Labour and Welfare Administration (NAV); the WP3 [1] of the MI-NAV project (ClinicalTrials.gov Identifier: NCT04196634). In the current study, all included participants will be pooled into one sample.

Study population, recruitment, stratification, randomization, and interventions

Eligible participants were workers, aged 18 to 67 years, on sick leave (50-100% sick leave rate, ≥ 7 consecutive weeks) due to musculoskeletal disorders (diagnosis within the musculoskeletal (L) chapter of the International Classification of Primary Care, 2nd edition (ICPC-2) [2]). Exclusion criteria were serious somatic or mental health disorders affecting work ability and entailing specialised treatment (e.g., cancer, psychotic disorders), pregnancy, unemployed, freelancers and self-employed workers, and insufficient Norwegian or English language skills to answer questionnaires or communicate by telephone. Participants were individually recruited through a phone call from NAV between April 2019 and October 2020. Every week recruiters received a list of eligible participants affiliated to eight NAV officers in South-Eastern Norway. All included participants signed an electronic informed consent form before study enrolment and were informed that they could withdraw at any time.

The Örebro Musculoskeletal Pain Screening Questionnaire Short Form (ÖMPSQ-SF) [3] and the Keele STarT MSK Tool (STarT MSK) [4] were used to stratify participants into risk groups of long-term sick leave [1, 5]. After stratification, participants were randomly allocated to either usual case management (UC), motivational interviewing (MI) and UC, or a stratified vocational advice intervention (SVAI) and UC, with a 1:1:1 allocation within each stratum of low/medium and high-risk.

A detailed description of the rationale, development and content of the interventions can be found elsewhere [1, 5]. Briefly, all participants were offered UC for people on sick leave in Norway. In addition, participants in the MI arm were offered 2 face-to-face sessions of MI from a NAV caseworker. Participants in the SVAI arm were offered vocational advice and case management from physiotherapists. Those stratified to the low/medium-risk group were offered 1-2 telephone sessions. Participants in the high-risk group were offered 3-4 sessions.

Data collection, outcome, modifiable prognostic factors, and covariates

At baseline, all participants responded to an electronic questionnaire including demographic variables and a set of patient-reported measures. Data on healthcare utilization were collected from public records including the Norwegian Patient Registry (NPR) and the Municipal Patient and User Registry (KPR). Data on productivity loss were collected from public records (NAV), containing dates and grading of absenteeism, work assessment allowance, and disability pension, as well as the related diagnostic code, and contracted workhours. Data on healthcare utilization and productivity loss were collected in the period from baseline to 3 months retrospectively, and in the 6 months follow-up period. To assess representativeness of the study sample, we obtained anonymised registry data covering sex, age, occupation, and contracted work hours from all eligible candidates. All information

is stored and will be analysed securely through the Service for sensitive data (TSD) at the University of Oslo, Norway.

Outcomes

The primary outcome of this study is costs related to healthcare utilization and productivity loss aggregated for 6 months of follow-up and dichotomized as high or low. Having high costs is defined as patients with costs in the top 25th percentile [6, 7]. As noted above, healthcare utilization was collected from public records (NPR, KPR) and included: primary healthcare use (general practitioner (GP), physiotherapist, chiropractor, and emergency room consultations) and secondary/tertiary healthcare use (outpatient contacts, day surgery, ordinary admission with overnight stay, and other admissions without overnight stay). Total cost of healthcare utilization per person will be estimated based on reimbursement rates collected from NPR and KPR. Productivity loss was collected from public records (NAV) and included productivity loss related to absenteeism, work assessment allowance, and disability pension. Total cost of productivity loss per person will be estimated based on number of days with productivity loss, adjusted for employment rate and grading of productivity loss, multiplied by an estimated average wage rate (from official statistics in Norway) including taxes and social costs. Healthcare utilization and productivity loss during the 6 months of follow-up will be described as shown in Table 2.

Secondary outcomes of this study are costs related to separately 1) healthcare utilization aggregated for 6 months of follow up and dichotomized as high and low, and 2) productivity loss aggregated for 6 months of follow up and dichotomized as high and low.

Modifiable prognostic factors

Potential modifiable prognostic factors are factors expected to have the potential to be modified or improved by appropriate care or treatment, and therefore classified as modifiable. Potential modifiable prognostic factors of high costs related to healthcare utilization and productivity loss includes the following self-reported variables measured at baseline:

- Pain severity measured by the numeric rating scale (NRS) [8] from the STarT MSK [4]
- Disability measured by a single item (Q3) from the EuroQol 5 dimensions (EQ-5D-5L) [9]
- Self-perceived health measured by a single item (Q6) from the EQ-5D-5L [9]
- Depressive symptoms measured by a single item (Q6) from the ÖMPSQ-SF [3]
- Sleep quality measured by a single item (Q4) from the ÖMPSQ-SF [3]
- Health literacy measured by a single item (Q12) from the Musculoskeletal Health Questionnaire (MSK-HQ) [10]
- Work satisfaction measured by a single item (0-10, 0 = not satisfied, 10 = satisfied)
- Long-lasting disorder expectation measured by a single item (Q6) from the STarT MSK [4]
- Return to work expectancy measured by a single item (Q8) from ÖMPSQ-SF [3]

Covariates

Prognostic factor research may vary depending on context (time, place, healthcare setting) and characteristics of the study population. We therefore plan to adjust for potential covariates when evaluating the modifiable prognostic factors. Potential covariates include the following self-reported variables measured at baseline:

- Sex
- Age
- Education level measured as the highest education completed, and categorised into low vs. high (university level)
- Pain duration measured by a single item (Q1) from the ÖMPSQ-SF [3]
- Group allocation (UC, MI and UC or SVAI and UC)

In addition, the following public records variables will be included as covariates:

- Absenteeism-related diagnosis type at baseline collected from the NAV registry and categorized into “upper/lower limb conditions”, “back and neck conditions”, “joint/inflammatory conditions”, “injury or trauma” or “other MSK conditions”
- Total costs related to healthcare utilization during a period of 3 months prior to inclusion. Healthcare utilization prior to inclusion was collected from public records (NPR, KPR) as described above. Total costs of healthcare utilization will be estimated as described above
- Total costs related to productivity loss during a period of 3 months prior to inclusion [11, 12]. Productivity loss prior to inclusion was collected from public records (NAV) as described above. Total costs of productivity loss will be estimated as described above

Other variables

Included participants will also be described with respect to the following baseline characteristics: mother tongue, days of productivity loss prior to inclusion, and healthcare utilization prior to inclusion.

Sample size

This study contains secondary analyses embedded in the MI-NAV project. Details on sample size calculation are provided in the MI-NAV protocol [1]. To determine statistical power of this study we used number of events per parameter (EPP) [13-17] and the rule-of-thumb of “10 events per parameter included” [18-21]. With a fixed sample size of 509 participants included in the MI-NAV project, we anticipate 127 participants to be in the top 25th percentile of costs and categorised as having high costs (yes/no) (events). An EPP of 10 will allow a maximum of 13 parameters to be included in the final multivariable prediction model.

Statistical analyses

General analysis considerations

All analyses described in this SAP are considered a priori in that they have been defined in the protocol and/or in this SAP. All post hoc analyses will be identified as such in the article if relevant. All analyses will be carried out using SPSS, Stata, R, or other appropriate software, and controlled by a senior researcher/statistician. We consider our study as explanatory. Thus, no correction for multiple testing will be performed and p-values < 0.05 will be considered statistically significant. All statistical tests will be two-sided. All confidence intervals will be reported as 95%. Preliminary analyses assessing the influence of missing data and assumptions of normality for continuous variables will be conducted. The assumption of normal distribution will be investigated using histograms and QQ-plots. Normally distributed data will be presented with means and standard deviations (SDs), skewed data with medians and interquartile range (IQR). Categorical data will be reported as counts and percentages.

Description of study flow

The flow of participants through the study will be reported with a flow chart according to the REMARK guidelines [22]. Reasons for dropout will be provided where known. Differences between responders and non-responders will be evaluated.

Missing data

We anticipate few missing values within this study. Information on the primary and secondary outcome will be obtained from public records (NPR, KPR, NAV) where all individuals receiving any form of benefits are registered by their social security number. Furthermore, we anticipate few missing data for most of the potential modifiable prognostic factors and the covariates, as a requirement to answer all questions on key questionnaires was implemented in the electronic baseline questionnaire. Nevertheless, missing value pattern will be visually explored and handled by multiple imputation if relevant (if >5% data is missing).

Participant characteristics

Baseline characteristics of included participants will be presented as shown in Table 1.

Healthcare utilization, productivity loss and cost estimation

Type and frequency of use of different healthcare resources will be calculated for the 6-month follow-up period. Costs related to healthcare utilization will be estimated based on reimbursement rates collected from NPR and KPR. Days of productivity loss will be calculated for the 6-month follow-up period and adjusted for employment rate and grading of productivity loss. Costs related to productivity loss will be estimated based on number of days with 100% productivity loss and national average wage rates (from official statistics) in Norway including taxes and social costs. Healthcare utilization and productivity loss will be presented as shown in Table 2. All costs will be presented in euros (€) 2022 and estimated with both mean and median values with 95%CI, using bias-corrected and accelerated (BCa) bootstrapping as presented in Table 3. The BCa will be conducted with a bootstrap sample size of 1000. Cost data are commonly skewed, thus both mean and median values will be presented to support the result interpretation. Values in Norwegian kroner (NOK) will be recalculated to euros using the exchange rate from January 2022 (1€=NOK 10).

External validation analysis

Univariable and multivariable binary logistic regression models will be used to external validate findings from the analysis within the MI-NAV WP2 material. Associations (crude and adjusted for selected covariates) between each predefined modifiable prognostic factor and costs related to 1) healthcare utilization and productivity loss, 2) healthcare utilization, and 3) productivity loss will be assessed. The cost score will be entered into the model as a dependent dichotomous variable (high cost defined as patients with cost in the top 25th percentile, yes/no). Non-linear relationships in the modelling process will be explored using cubic splines or multivariable fraction polynomials, as these are recommended approaches for modelling continuous prognostic factors in prognosis research [23]. The results will be presented as crude and adjusted odds ratios with 95% confidence intervals as shown in Table 4. The decision on whether findings from WP2 are replicated will be based on the size and direction of the association, the confidence interval, and the p-value for each of the predefined prognostic factors [24].

Sensitivity analysis

To assess credibility of the total cost calculation related to the primary analyses, the calculation will be conducted without outliers. Outliers will be identified with simple scatterplots by visual inspection and defined as patients with remarkably high total costs. If multiple imputation on missing data is conducted, the univariable and multivariable logistic regression analyses related to the primary analyses will be performed on complete case data to test credibility of the imputation procedure.

Ethics approval

This study is a part of the MI-NAV project [1, 25]. The MI-NAV project (ClinicalTrials.gov Identifier: NCT04196634) has been classified as a quality assessment study by the Norwegian Regional Committee for Medical Research Ethics (reference no. 2018/1326/REK sør-øst A) and approved by the Norwegian Social Science Data Service (reference no. 861249) in 2018.

Table 1. Participants characteristics and clinical status at baseline

	All participants (n=)	Missing, n (%)
Female, n (%)		
Age in years		
Education at university level, n (%)		
Mother tongue Norwegian, n (%)		
Diagnosis (ICPC-2)*, n (%)		
Upper limb conditions		
Lower limb conditions		
Neck conditions		
Back conditions		
Joint or inflammatory conditions		
Injurie or trauma		
Other MSK conditions		
Pain severity average last week (NRS, 0-10)		
Pain duration, n (%)		
< 3 months		
3-6 months		
> 6 months		
Disability (EQ-5D-5L, Q3), n (%)		
No problems doing usual activities		
Slight problems doing usual activities		
Moderate problems doing usual activities		
Severe problems doing usual activities		
Unable to do usual activities		
Self-perceived health (EQ-5D-5L, Q6, 0-10)		
Depressive symptoms (ÖMPSQ-SF, Q6, 0-10)		
Sleep quality (ÖMPSQ-SF, Q4, 0-10)		
Health literacy (MSK-HQ, Q12), n (%)		
Completely understanding of condition/treatment		
Very well understanding of condition/treatment		
Moderately understanding of condition/treatment		
Slightly understanding of condition/treatment		
No understanding of condition/treatment		
Long-lasting disorder expectation (STarT MSK, Q6), n (%)		
Return to work expectancy (ÖMPSQ-SF, Q8, 0-10)		
Work satisfaction (0-10)		
*Healthcare utilization prior to inclusion***		
Primary care consultation last 3 months, n (%)		
General practitioner		
Physiotherapist		
Chiropractor		
Emergency room		
Secondary/tertiary care last 3 months, n (%)		
Outpatient contact		
Day surgery		
Ordinary admission with overnight stay		
Other admissions without overnight stay		
*Productivity loss prior to inclusion****		
Days of sick leave last 3 months		
Days of work assessment allowance last 3 months		
Days of disability benefits last 3 months		

EQ-5D-5L indicates EuroQol 5 dimensions; ICPC-2, International Classification of Primary Care 2ed edition; MSK-HQ, Musculoskeletal Health Questionnaire; NRS, Numeric Rating Scale; ÖMPSQ-SF, Örebro Musculoskeletal Pain Screening Questionnaire Short Form; STarT MSK, Keele STarT MSK tool. *Absenteeism related diagnoses type at baseline, collected from the Norwegian Labour and Welfare Administration (NAV) registry. **Collected from public records; the Norwegian Patient Registry (NPR) and the Municipal Patient and User Registry (KPR). ***Collected from the NAV registry, measured as calendar days, and adjusted for employment rate and grading of productivity loss.

Table 2. Healthcare utilization and productivity loss throughout 6-month of follow-up

	All participants (n=)	Missing, n (%)
Primary care		
Participants with primary care consultation, n (%)		
General practitioner		
Physiotherapist		
Chiropractor		
Emergency room		
No primary care consultation		
Numbers of consultations, median (IQR)*		
General practitioner		
Physiotherapist		
Chiropractor		
Emergency room		
Secondary/tertiary care		
Participants with secondary/tertiary care consultation, n (%)		
Outpatient contact		
Day surgery		
Ordinary admission with overnight stay		
Other admissions without overnight stay		
No secondary/tertiary care consultation		
Numbers of consultations, median (IQR)*		
Outpatient contact		
Day surgery		
Ordinary admission with overnight stay		
Other admissions without overnight stay		
Duration of ordinary admission with overnight stay in days, median (IQR)**		
Productivity loss		
Participants with productivity loss, n (%)		
Sick leave		
Work assessment allowance		
Disability benefits		
Duration of productivity loss in days, median (IQR)***		
Sick leave		
Work assessment allowance		
Disability benefits		

*Calculated on basis of participants who have reported primary/secondary/tertiary care consultations. **Calculated on basis of participants who have reported ordinary admission with overnight stay. ***Calculated on basis of participants who have reported productivity loss, converted into a 5-day workweek, and adjusted for employment rate and grading of productivity loss.

Table 3. Cost (€) due to healthcare utilization and productivity loss throughout 6-month of follow-up

	Participants with zero cost, n (%)	Mean (95%CI*)	Median (95% CI*)
Primary care			
General practitioner			
Physiotherapist			
Chiropractor			
Emergency room			
Total			
Secondary/tertiary care			
Outpatient contact			
Day surgery			
Ordinary admission with overnight stay			
Other admissions without overnight stay			
Total			
Productivity loss			
Sick leave			
Work assessment allowance			
Disability benefits			
Total			
Total			

Cost related to productivity loss are calculated on basis of reported days with productivity loss, converted into a 5-day workweek, and adjusted for employment rate and grading of productivity loss. *Bias-corrected and accelerated bootstrapping (1000 simulations).

Table 4. Binary logistic regression analyses; individual associations between modifiable prognostic factors and high costs

	High costs related to healthcare utilization and productivity loss		**High costs related to healthcare utilization**		**High costs related to productivity loss**	
	Crude OR (95% CI)	Adjusted OR* (95% CI)	Crude OR (95% CI)	Adjusted OR* (95% CI)	Crude OR (95% CI)	Adjusted OR* (95% CI)
Pain severity (NRS, 0-10)						
Self-perceived health (EQ-5D-5L, Q6, 0-10)						
Depressive symptoms (ÖMPSQ-SF, Q6, 0-10)						
Sleep quality (ÖMPSQ-SF, Q4, 0-10)						
Return to work expectancy (ÖMPSQ-SF, Q8, 0-10)						
Work satisfaction (0-10)						
Disability (EQ-5D-5L, Q3) (ref: no/slight problems)						
Moderate problems						
Severe problems/unable to do						
Health literacy (MSK-HQ, Q12) (ref: completely/very well understanding)						
Moderate understanding						
Slightly/no understanding						
Long-lasting disorder expectation (STarT MSK, Q6) (ref: no)						

EQ-5D-5L indicates EuroQol 5 dimensions; CI, confidence interval; MSK-HQ, Musculoskeletal Health Questionnaire; NRS, Numeric Rating Scale; ÖMPSQ-SF, Örebro Musculoskeletal Pain Screening Questionnaire Short Form; STarT MSK, Keele STarT MSK tool; Q, question number. *Adjusted by sex, age, education level, absenteeism related diagnosis type, pain duration, and costs related to 1) healthcare utilization and productivity loss prior to inclusion, 2) healthcare utilization prior to inclusion, or 3) productivity loss prior to inclusion.

References

1. Oiestad, B.E., et al., *Study protocol for a randomized controlled trial of the effectiveness of adding motivational interviewing or stratified vocational advice intervention to usual case management on return to work for people with musculoskeletal disorders. The MI-NAV study.* BMC Musculoskelet Disord, 2020. **21**(1): p. 496.
2. Classification Committee WONCA, *ICPC-2: International Classification of Primary Care*. 2nd ed. 1998, Oxford: Oxford University Press.
3. Linton, S.J., M. Nicholas, and S. MacDonald, *Development of a short form of the Orebro Musculoskeletal Pain Screening Questionnaire.* Spine (Phila Pa 1976), 2011. **36**(22): p. 1891-5.
4. Hill, J.C., et al., *Does a modified STarT Back Tool predict outcome with a broader group of musculoskeletal patients than back pain? A secondary analysis of cohort data.* BMJ Open, 2016. **6**(10): p. e012445.
5. Aanesen, F., et al., *Effectiveness of adding motivational interviewing or a stratified vocational advice intervention to usual case management on return to work for people with musculoskeletal disorders: the MI-NAV randomised controlled trial.* Occup Environ Med, 2023. **80**(1): p. 42-50.
6. Engel, C.C., M. von Korff, and W.J. Katon, *Back pain in primary care: predictors of high health-care costs.* Pain, 1996. **65**(2-3): p. 197-204.
7. Becker, A., et al., *Low back pain in primary care: costs of care and prediction of future health care utilization.* Spine (Phila Pa 1976), 2010. **35**(18): p. 1714-20.
8. Von Korff, M., M.P. Jensen, and P. Karoly, *Assessing global pain severity by self-report in clinical and health services research.* Spine (Phila Pa 1976), 2000. **25**(24): p. 3140-51.
9. Herdman, M., et al., *Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L).* Qual Life Res, 2011. **20**(10): p. 1727-36.
10. Hill, J.C., et al., *Development and initial cohort validation of the Arthritis Research UK Musculoskeletal Health Questionnaire (MSK-HQ) for use across musculoskeletal care pathways.* BMJ Open, 2016. **6**(8): p. e012331.
11. Lentz, T.A., et al., *Factors associated with persistently high-cost health care utilization for musculoskeletal pain.* PLoS One, 2019. **14**(11): p. e0225125.
12. Budtz, C.R., S. Mose, and D.H. Christiansen, *Socio-demographic, clinical and psychological predictors of healthcare utilization among patients with musculoskeletal disorders: a prospective cohort study.* BMC Health Serv Res, 2020. **20**(1): p. 239.
13. Harrell, F.E., Jr., et al., *Regression modelling strategies for improved prognostic prediction.* Stat Med, 1984. **3**(2): p. 143-52.
14. Harrell, F.E., Jr., K.L. Lee, and D.B. Mark, *Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors.* Stat Med, 1996. **15**(4): p. 361-87.
15. Steyerberg, E.W., et al., *Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets.* Med Decis Making, 2001. **21**(1): p. 45-56.
16. Steyerberg, E.W., et al., *Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets.* Stat Med, 2000. **19**(8): p. 1059-79.
17. Ambler, G., A.R. Brady, and P. Royston, *Simplifying a prognostic model: a simulation study based on clinical data.* Stat Med, 2002. **21**(24): p. 3803-22.
18. Moons, K.G., et al., *Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist.* PLoS Med, 2014. **11**(10): p. e1001744.
19. Pavlou, M., et al., *Review and evaluation of penalised regression methods for risk prediction in low-dimensional data with few events.* Stat Med, 2016. **35**(7): p. 1159-77.
20. Pavlou, M., et al., *How to develop a more accurate risk prediction model when there are few events.* BMJ, 2015. **351**: p. h3868.
21. Cowley, L.E., et al., *Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature.* Diagn Progn Res, 2019. **3**: p. 16.
22. McShane, L.M., et al., *REporting recommendations for tumour MARKer prognostic studies (REMARK).* Br J Cancer, 2005. **93**(4): p. 387-91.
23. Riley, R.D., et al., *Prognosis Research in Healthcare. Concepts, Methods, and Impact*. First ed. 2019, Oxford: Oxford United Press.
24. Kent, P., et al., *A conceptual framework for prognostic research.* BMC Med Res Methodol, 2020. **20**(1): p. 172.
25. Tveter, A.T., et al., *Risk assessment for prolonged sickness absence due to musculoskeletal disorders: protocol for a prospective cohort study.* BMC Musculoskelet Disord, 2020. **21**(1): p. 326.