Statistical analysis plan (SAP) for the project:
BACk pain in Elders in Norway (BACE-N): a prospective cohort study of older people visiting primary care with a new episode of back pain

SAP Authors:
Rikke Munk Killingmo, Ørjan Nesse Vigdal, Lise Kretz, Kjersti Storheim, Are Hugo Pripp, Margreth Grotle (order not decided)

NCT identifier: NCT04261309
Document date: 21.02.2024

Statistical analysis plan (SAP) for:

Healthcare utilization and related costs among older people seeking primary care due to back pain: a prospective cohort study with one year of follow up (working title)

Project:	BACk pain in Elders in Norway (BACE-N) (BACE-N)
NCT number:	NCT04261309
Document date:	11.03.2021

1. Introduction to SAP

1.1 Scope

This document is a supplement to the BACE-N protocol (ClinicalTrials.gov Identifier: NCT04261309) and comprises a SAP for the article "Healthcare utilization and related cost among older people seeking primary care due to back pain: a prospective cohort study with one year of follow up". The current SAP has been written while data collection was ongoing (we had access to baseline data, but not to follow-up data) and it will be uploaded to the ClinicalTrials.gov before full access to the study database.

2. Administrative information

Version of SAP

1.0

Study sponsor

Oslo Metropolitan University, the Norwegian Fund for Post-Graduate Training in Physiotherapy and "Et liv i bevegelse" (A life in movement) - Norwegian chiropractors' research foundation

Names, affiliations and roles of SAP contributors

Name	Title	Affiliation	Role
Margreth Grotle	Professor	Oslo Metropolitan University P44, 0167 Oslo Phone: 90111172 Email: mgrotle@oslomet.no	Principal Investigator
Rikke Munk Killingmo	PhD-student	Oslo Metropolitan University	Main author of SAP
Kjersti Storheim	Professor	Oslo Metropolitan University	Contributor to SAP
Daniëlle van der Windt	Professor	Keele University	Contributor to SAP
Milada C. Småstuen	Professor	Oslo Metropolitan University	Statistical advisor
Ørjan Nesse Vigdal	PhD-student	Oslo Metropolitan University	Contributor to SAP
Lise Kretz	PhD-student	Oslo Metropolitan University	Contributor to SAP

Signatures: person writing SAP, senior statistician and principal investigator

Rikke Munk Killingmo	**Milada C. Småstuen**	**Margreth Grotle**
Main author of the SAP	Senior statistician responsible	Principal Investigator

3. Study aim

The primary aim of this study is to describe healthcare utilization and estimate associated costs during one year of follow up of older people seeking primary care due to a new episode of back pain. The secondary aim is to describe frequency of healthcare utilization across patients with different risk profile stratified according to the StarT Back Screening tool.

4. Study design, population and method

Study design

A prospective observational cohort study with one year of follow-up within a Norwegian primary care setting.

Study population and recruitment

Eligible patients are people 55 years of age or older who seek primary care (physiotherapist, chiropractor or GP) with a new episode of back pain (preceded by 6 months without visiting a primary care provider for similar complaints). Patients are excluded if they have difficulties completing the questionnaires (e.g. unable to speak, read or write in Norwegian) or if they have difficulties completing the physical examination (e.g. are wheelchair bound).

Patients were recruited from physiotherapists, chiropractors and GPs working in Norwegian primary care between April 2015 and February 2020. Patients who met the eligibility criteria and completed the consent to participate were included in the study.

Method

At baseline all patients completed a comprehensive questionnaire and underwent a standardized physical examination conducted by local research assistants at test stations established within each recruiting area. Follow-up questionnaires will be sent at 3, 6, and 12 months after inclusion for completion at home. All questionnaires are preferably completed electronically using the Infopad system, but paper versions will be available for patients not familiar with electronic data collection. All information will be stored and analysed securely through Service for sensitive data (TSD) at the University of Oslo, Norway.

Variables

Outcome variables

The primary outcome of this study is total costs of healthcare utilization (summed up for one year of follow-up). The secondary outcome is frequency of healthcare utilization (summed up for one-year of follow-up across patients with different risk profile according to the StarT Back Screening tool).

Healthcare utilization will be self-reported, collected through follow-up questionnaires, and include; consultation to healthcare professionals (type and frequency), number of diagnostic examinations (type and frequency), number of days of hospitalization and/or institutionalisation, back operations and use of back medication (both prescription and over-the-counter, type and frequency). All variables, except the “operation” variable, will be

reported with a 3 month recall period at 3, 6, and 12 months follow-up. The "operation" variable will be reported with a 12 month recall period at 12 months follow-up. Healthcare utilization during the one year of follow-up will be described as shown in table 3 and table 5. The total cost of healthcare utilization will be estimated based on information presented in table 3 and unit costs of healthcare resources collected from national pricelists (see table 1).

Screening tool

The Keele StarT Back Screening Tool (SBST) [1] will be used to classify the included patients into low, medium or high risk of poor disability outcome. The SBST is a brief 9-item tool designed to screen primary care patients with low back pain for prognostic indicators that are relevant to initial decision making. The tool is summed to produce an overall score from zero to nine and a psychological subscale score from zero to five. Patients with an overall score between 0-3 are classified as low risk. Patients with an overall score of minimum 4 and a subscale score of maximum 3 are classified as medium risk. Patient with an overall score of minimum 4 and a subscale score of 4 or 5 are classified as high risk. The SBST was translated into Norwegian by Storheim and Grotle in 2012, and has shown to have an acceptable accuracy in distinguishing between low back pain patients who have recovered or not after 1 year of follow-up [2].

Other variables

Included patients will be described with respect to the following baseline characteristics: age, gender, ethnicity, educational level, first healthcare provider, pain location, pain severity, pain duration, pain history, disability, comorbidity, health-related quality of life, emotional well-being, kinesiophobia and healthcare utilization six weeks prior to inclusion (table 2).

The Numeric Rating Scale (NRS) will be used to measure average pain severity last week [3]. The NRS, scored from 0 (no pain) to 10 (maximum pain), has been widely used to evaluate pain and has proven to be preferable when examining low back pain patients [4], also for Norwegian patients [5].

The Roland Morris Disability Questionnaire (RMDQ) [6] will be used to measure disability. The RMDQ is a widely used back-specific patient-reported measure of pain-related disability (0 = no disability, 24 = totally disabled). The Norwegian version has been validated and found to have good measurement properties when used among patients with low back pain [5, 7].

The Self-Administered Comorbidity Questionnaire (SCQ) [8] will be used to assess comorbidity. The SCQ is a 14-item measure of comorbidity for clinical and health services research settings. An individual can receive a maximum of 3 points for each medical condition: 1 point for the presence of the problem, another point if he/she receives treatment for it, and an additional point if the problem causes a limitation in functioning. Because there are 12 defined medical problems and 3 optional conditions, the maximum score totals 45 points if the open-ended items are used and 36 points if only the close-ended items are used.

The Short Form-36 Health Status Questionnaire (SF36) [9] will be used to assess health-related quality of life. The SF36 consist of 36 items. It measures health on eight multi-item

dimensions, covering physical functioning, social functioning, role limitations (physical problems), role limitations (emotional problems), mental health, vitality, pain, and overall evaluation of health [9]. Data-completeness of the SF36 in the general population in Norway seems to strongly declined with increasing age [10]. Hence, caution should be exercised when assessing subjective health or employing the norms among subjects aged 70 years or over [10].

The Center for Epidemiologic Studies Depression Scale (CES-D) will be used to assess emotional well-being. The CES-D has been widely used in studies of late-life depression. Psychometric properties are generally favourable [11]. The Norwegian version of the CES-D has been used among older patients in order to measure depression symptoms [12].

The Fear Avoidance Beliefs Questionnaire, physical activity subscale (FABQ-PA) [13] will be used to assess kinesiophobia. The FABQ-PA consists of four questions aimed towards physical activity, scored on a 7-point ordinal scale, which are summed up to a sum score, ranging from 0 (no fear) to 24 (maximum fear). The questionnaire has been translated into Norwegian and has shown acceptable psychometric properties in Norwegian patients with low back pain [14].

5. Statistical analyses

General analysis considerations

All analyses described in this plan 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 by a PhD-student using SPSS version 26 and controlled by a senior researcher/statistician. All statistical tests will be two-sided, and nominal p-values will be reported. 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. Missing data will be handled by multiple imputation, using 5 imputations and 10 iterations unless the missingness exceeds 30% and/or missing at random cannot be assumed. Fully conditioned specification method and regression estimation will be used. For variables where we are unable to use regression estimation due to computational difficulties, predictive mean matching will be used [15].

Description of study flow

The flow of participants through the study will be reported according to the STROBE guidelines [16] with a flow chart (see figure 1). Reasons for dropout will be provided where known. Differences between responders and non-responders will be evaluated.

Participant characteristics

Baseline characteristics of included patients will be presented for the whole sample and the following subgroups: (1) low, (2) medium, and (3) high risk of persistent disabling back pain according to the StarT Back Screening tool (see table 2).

Primary analysis

First, type and frequency of different healthcare resources will be calculated for each of the follow-up periods; from baseline to 3 months, 3 to 6 months, and 9 to 12 months and presented as shown in table 3.

Next, costs will be estimated based on information presented in table 3 and unit costs of healthcare resources collected from national pricelists (see table 1). Costs related to back medication will be estimated based on medication type (not exact medication name) and frequency of use. Data on dosage is not available. All costs will be presented in Norwegian Krones (NOK) and euros (€) 2020. Costs of healthcare utilization will be described with median and interquartile range for each follow-up period and for the whole year as shown in table 4. If data is highly skewed, bias-corrected and accelerated bootstrapping will be considered to derive confidence intervals for cost estimates.

Secondary analysis

First, type and frequency of different healthcare resources will be described for the one-year of follow-up for the following subgroups: (1) low, (2) medium, and (3) high risk of persistent disabling back pain according to the StarT Back Screening tool and presented as shown in table 5.

Next, Kruscall Wallis test will be used to determine whether there is a significant difference between the three subgroups with regards to; number of primary care consultations, number of patients using back medication, number of patients receiving imaging (X-ray, MRI, CT) and number of patients receiving secondary care (back operation, hospitalization, rehabilitation stay). For all analyses, p values of < 0.05 will be considered statistically significant.

Sample size and statistical power consideration

This study contains secondary analyses embedded in the BACE-N study. Details on sample size calculation are provided in the BACE-N protocol (ClinicalTrials.gov Identifier: NCT04261309). With a sample size of 450 participants within the BACE-N study, we assume to have expectable power to describe healthcare utilization and estimate associated costs [17].

Sensitivity analyses

To assess the robustness of the results, the following sensitivity analyses will be carried out for cost calculations related to the primary analysis:

1. Complete case analysis (without using imputation for missing data)
2. Without outliers (outliers will be identified with simple scatter plots by visual inspection)

6. Selection bias, information bias and confounding

Selection bias:

Because of limited resources and practical reasons related to recruitment from a broad network of clinicians, the BACE-N lacks information on eligible study participants that declined to participate or for other reasons were not invited. Therefore, in order to assess representativeness, the BACE-N study sample will be compared on key sociodemographic variables with a sample from the longitudinal population study of people in the second half of life; The Norwegian study on life course, ageing and generation (NORLAG) [18]. The NORLAG study is expected to represent a representative sample of older people with musculoskeletal complaints.

Response rate at each assessment point and reasons for loss to follow-up will be reported. Key baseline characteristics will be compared between those lost to follow-up and those remaining in the study.

Information bias:

To reduce the risk of information bias, the study outcome (costs) will be measured in an identical manner in all included cases, and in the best possible way within the framework of the BACE-N study.

Covariates:

Covariates may influence estimates of our primary outcome. Therefore, in line with the PROGRESS framework and recommendations for type 1 studies [19], we will describe potential covariates and the way they may alter our primary outcome.

Potential covariates of costs related to healthcare utilization: gender [20-25], age [20, 22, 26], educational level [27, 28], pain duration [21, 28-30], pain history [25], pain severity [22, 25, 30, 31], radiating pain below the knee [29], disability [21, 23, 25, 28, 29], comorbidity [26, 32, 33], health-related quality of life [23], emotional well-being [21, 28, 29, 31, 32], kinesiophobia [29, 34], first healthcare provider [35] and costs related to healthcare utilization prior to inclusion.

Table 1 Cost categories, units, unit price, all numbers in Euros (€) and Norwegian kroner (NOK) for 2020

Cost categories	Unit	Unit price (€)	Unit price (NOK)	Reference (source)
Primary care				
General practitioner	Per visit			
Physiotherapist	Per visit			
Chiropractor	Per visit			
Manuel therapist	Per visit			
Naprapath	Per visit			
Osteopath	Per visit			
Psychologist	Per visit			
Other therapists	Per visit			
Back medication				
Paracetamol	Per daily defined dose			
NSAID	Per daily defined dose			
Muscle relaxant	Per daily defined dose			
Sleep medication	Per daily defined dose			
Cortisone	Per daily defined dose			
Opioid	Per daily defined dose			
Others	Per daily defined dose			
Examinations				
Blood sample	Per examination			
X-ray	Per examination			
MRI	Per examination			
CT	Per examination			
Others??	Per examination			
Secondary care				
Back operation	Per operation			
Hospitalization (non-operation)	Per day			
Rehabilitation stay	Per day			

NoMA, Norwegian Medicines Agency

Table 2 Patient characteristics and clinical status at baseline (n = X)

	All participants (n = x)	Stratified risk profile*		
		Low (n = x)	Medium (n = x)	High (n = x)
Male, N (%)				
Age in years, mean (SD)				
Educational level high, N (%)				
Ethnicity Norwegian, N (%)				
First healthcare provider, N (%)				
General practitioner				
Physiotherapist				
Chiropractor				
Pain location, N (%)				
Lumbar				
Thoracic				
Radiating pain below the knee				
Pain severity last week (NRS 0-10), median (IQR)				
Pain duration, N (%)				
< 6 weeks				
6 weeks to 3 months				
> 3 months				
Previous episodes of back pain, N (%)				
Disability (RMDQ 0-24), mean (SD)				
Comorbidity (SCQ, 0-15)				
Health-related QOL (SF36, 0-100), mean (SD)				
Physical component				
Mental component				
Emotional well-being (CES-D 0-60)				
Kinesiophobia (FABQ-PA 0-24)				
Healthcare utilization prior to inclusion				
Patients with primary care consultation last 6 weeks, N (%)				
General practitioner				
Physiotherapist				
Chiropractor				
Manual therapist				
Naprapath				
Osteopath				
Psychologist				
Other therapists				
Patients with use of back medication, N (%)?				
Patients with additional diagnostic examination last 6 months, N (%)				
Blood sample				
X-ray				
MRI				
CT				
Patients with previous hospitalization, N (%)				
Patients with previous rehabilitation stay, N (%)				

* According to the StarT Back Screening Tool

Table 3 Healthcare utilization throughout one-year of follow-up

	0-3 months	>3-6 months	>9-12 months
Primary care			
Patients with primary care consultation, N (%)			
General practitioner			
Physiotherapist			
Chiropractor			
Manual therapist			
Naprapath			
Osteopath			
Psychologist			
Other therapists			
Number of general practitioner consultations, median (IQR)			
Number of physiotherapist consultations, median (IQR)			
Number of chiropractor consultations, median (IQR)			
Number of manual therapist consultations, median (IQR)			
Number of naprapath consultations, median (IQR)			
Number of psychologist consultations, median (IQR)			
Number of other consultations, median (IQR)			
Back medication			
Patients with use of back medication, N (%)			
Paracetamol			
NSAID			
Muscle relaxants			
Sleep medication			
Cortisone			
Opioid			
Others			
Frequency of use paracetamol, N (%)			
Daily			
Weekly			
Monthly or less			
Frequency of use NSAID, cortisone N (%)			
Daily			
Weekly			
Monthly or less			
Frequency of use muscle relaxants, sleep medication, N (%)			
Daily			
Weekly			
Monthly or less			
Examinations			
Patients with additional diagnostic examination, N (%)			
Blood sample			
X-ray			
MRI			
CT			
Secondary care			
Patients with back operation, N (%)			
Patients with hospitalization, N (%)			
Duration of stay in days, median (IQR)			
Patients with rehabilitation stay, N (%)			
Duration of stay in days, median (IQR)			

Table 4 Cost due to healthcare utilization from 0-3 month, >3-6 months, >9-12 months and the entire follow-up period (0-12 month)

	0-3 months	3-6 months	>9-12 months	0-12 months*
Primary care				
General practitioner				
Physiotherapist				
Chiropractor				
Manual therapist				
Naprapath				
Osteopath				
Psychologist				
Other therapists				
Back medication				
Paracetamol				
NSAID, cortisone				
Muscle relaxants, sleep medication				
Examinations				
Blood sample				
X-ray				
MRI				
CT				
Secondary care				
Back operation				
Hospitalization and/or rehabilitation stay				
Total costs				

Values are median (interquartile range) of costs (€). *Cost due to healthcare utilization for the entire follow-up period is calculated on the basis for the three follow-up periods

Table 5 Healthcare utilization throughout one-year of follow-up, across patients with different risk profile according to the StarT Back Screening tool

	Stratified risk profile		
	Low (n = x)	Medium (n = x)	High (n = x)
Primary care			
Patients with primary care consultation, N (%)			
Number of primary consultations, median (IQR)			
Back medication			
Patients with use of back medication, N (%)			
Paracetamol			
NSAID			
Muscle relaxants			
Sleep medication			
Cortisone			
Opioid			
Examinations			
Patients with additional diagnostic examination, N (%)			
Blood sample			
X-ray			
MRI			
CT			
Secondary care			
Patients with back operation, N (%)			
Patients with hospitalization, N (%)			
Patients with rehabilitation stay, N (%)			

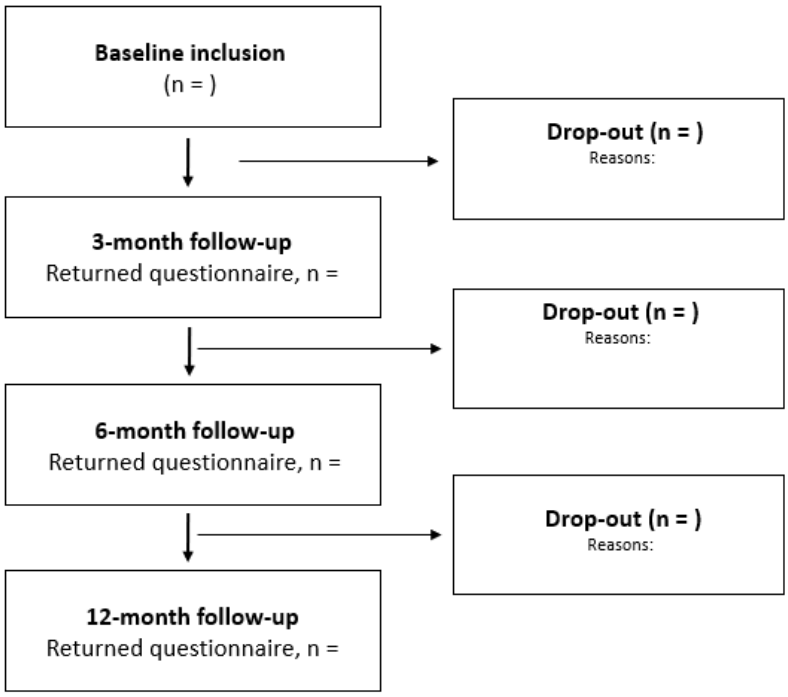


Figure 1. Flow chart of the study

References

1. Hill, J.C., et al., *A primary care back pain screening tool: identifying patient subgroups for initial treatment.* Arthritis Rheum, 2008. **59**(5): p. 632-41.
2. Vigdal Ø. N.; Storheim K., G.M., *With what accuracy can we predict persistent disability in sciatica patients with self-reported screening tools?* 2014, Oslo University College.
3. 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.
4. Strong, J., R. Ashton, and D. Chant, *Pain intensity measurement in chronic low back pain.* Clin J Pain, 1991. **7**(3): p. 209-18.
5. Grotle, M., J.I. Brox, and N.K. Vollestad, *Concurrent comparison of responsiveness in pain and functional status measurements used for patients with low back pain.* Spine (Phila Pa 1976), 2004. **29**(21): p. E492-501.
6. Roland, M. and R. Morris, *A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain.* Spine (Phila Pa 1976), 1983. **8**(2): p. 141-4.
7. Grotle, M., J.I. Brox, and N.K. Vollestad, *Cross-cultural adaptation of the Norwegian versions of the Roland-Morris Disability Questionnaire and the Oswestry Disability Index.* J Rehabil Med, 2003. **35**(5): p. 241-7.
8. Sangha, O., et al., *The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research.* Arthritis Rheum, 2003. **49**(2): p. 156-63.
9. Ware, J.E., Jr. and C.D. Sherbourne, *The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection.* Med Care, 1992. **30**(6): p. 473-83.
10. Loge, J.H. and S. Kaasa, *Short form 36 (SF-36) health survey: normative data from the general Norwegian population.* Scand J Soc Med, 1998. **26**(4): p. 250-8.
11. Radloff, L.S., *The CES-D scale: a Self-Report Depression Scale for Research in the General Population.* Applied psychological measurement, 1977. (1(3)): p. 385-355.
12. SE., T.K.C., *Funksjonshemning, ensomhet og depresjon: Hva betyr ensomhet for om personer med funksjonshemning opplever depresjon?* Tidsskrift Norsk Psykologforening, 2008. **1**.
13. Waddell, G., et al., *A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability.* Pain, 1993. **52**(2): p. 157-68.
14. Grotle, M., J.I. Brox, and N.K. Vollestad, *Reliability, validity and responsiveness of the fear-avoidance beliefs questionnaire: methodological aspects of the Norwegian version.* J Rehabil Med, 2006. **38**(6): p. 346-53.
15. Van Buuren, S., *Flexible imputation of missing data*. 2018: CRC press.
16. von Elm, E., et al., *The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies.* PLoS Med, 2007. **4**(10): p. e296.
17. Johnston, K.M., et al., *Methods of sample size calculation in descriptive retrospective burden of illness studies.* BMC Med Res Methodol, 2019. **19**(1): p. 9.
18. Slagsvold, B.V., M.; Daatland, S. O.; Hagestad, G.; Hansen, T.; Herlofson K.; Koløen, K; Solem, P. E. , *Life-course, ageing and generations in Norway:the NorLAG study.* Norsk Epidemiologi, 2012. **22**(2).
19. Hemingway, H., et al., *Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes.* BMJ, 2013. **346**: p. e5595.
20. Rattay, P., et al., *[Utilization of outpatient and inpatient health services in Germany: results of the German Health Interview and Examination Survey for Adults (DEGS1)].* Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz, 2013. **56**(5-6): p. 832-44.
21. Chenot, J.F., et al., *Sex differences in presentation, course, and management of low back pain in primary care.* Clin J Pain, 2008. **24**(7): p. 578-84.
22. Wenig, C.M., et al., *Costs of back pain in Germany.* Eur J Pain, 2009. **13**(3): p. 280-6.

23. Mutubuki, E.N., et al., *Predictive factors of high societal costs among chronic low back pain patients.* Eur J Pain, 2019.
24. Ekman, M., et al., *Burden of illness of chronic low back pain in Sweden: a cross-sectional, retrospective study in primary care setting.* Spine (Phila Pa 1976), 2005. **30**(15): p. 1777-85.
25. Ferreira, M., et al., *Factors defining care-seeking in low back pain - A meta-analysis of population based surveys.* European Journal of Pain, 2010. **14**(7): p. 747.e1-747.e7.
26. Lentz, T.A., et al., *Factors associated with persistently high-cost health care utilization for musculoskeletal pain.* PLoS One, 2019. **14**(11): p. e0225125.
27. Hoebel, J., et al., *Socioeconomic Status and Use of Outpatient Medical Care: The Case of Germany.* PLoS One, 2016. **11**(5): p. e0155982.
28. Lim, K.L., P. Jacobs, and S. Klarenbach, *A population-based analysis of healthcare utilization of persons with back disorders: results from the Canadian Community Health Survey 2000-2001.* Spine (Phila Pa 1976), 2006. **31**(2): p. 212-8.
29. 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.
30. 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.
31. Hirsch, O., et al., *Low back pain patient subgroups in primary care: pain characteristics, psychosocial determinants, and health care utilization.* Clin J Pain, 2014. **30**(12): p. 1023-32.
32. Stewart, W.F., et al., *Patterns of health care utilization for low back pain.* J Pain Res, 2015. **8**: p. 523-35.
33. Ritzwoller, D.P., et al., *The association of comorbidities, utilization and costs for patients identified with low back pain.* BMC Musculoskelet Disord, 2006. **7**: p. 72.
34. Keeley, P., et al., *Psychosocial predictors of health-related quality of life and health service utilisation in people with chronic low back pain.* Pain, 2008. **135**(1-2): p. 142-50.
35. Fritz, J.M., J. Kim, and J. Dorius, *Importance of the type of provider seen to begin health care for a new episode low back pain: associations with future utilization and costs.* J Eval Clin Pract, 2016. **22**(2): p. 247-52.

Statistical analysis plan (SAP) for:

Modifiable prognostic factors of high cost related to healthcare utilization among older people seeking primary care with a new episode of back pain - an identification and replication study (working title)

Project:	BACk pain in Elders in Norway (BACE-N) (BACE-N)
NCT number:	NCT04261309
Document date:	11.03.2021

1. Introduction to SAP

1.1 Scope

This document is a supplement to the BACE-N protocol (ClinicalTrials.gov Identifier: NCT04261309) and comprises a SAP for the article "Modifiable prognostic factors of high cost related to healthcare utilization among older people seeking primary care with a new episode of back pain - an identification and external validation study". The current SAP has been written while data collection was ongoing (we had access to baseline data, but not to follow-up data) and it will be uploaded to the ClinicalTrials.gov before full access to the study database.

2. Administrative information

Version of SAP

1.0

Study sponsor

Oslo Metropolitan University, the Norwegian Fund for Post-Graduate Training in Physiotherapy and "Et liv i bevegelse" (A life in movement) - Norwegian chiropractors' research foundation

Names, affiliations and roles of SAP contributors

Name	Title	Affiliation	Role
Margreth Grotle	Professor	Oslo Metropolitan University P44, 0167 Oslo Phone: 90111172 Email: mgrotle@oslomet.no	Principal Investigator
Rikke Munk Killingmo	PhD-student	Oslo Metropolitan University	Main author of SAP
Kjersti Storheim	Professor	Oslo Metropolitan University	Contributor to SAP
Daniëlle van der Windt	Professor	Keele University	Contributor to SAP
Bart Koes	Professor	Erasmus MC	Contributor to SAP
Sita Bierma-Zienstra	Professor	Erasmus MC	Contributor to SAP
Alessandro Chiarotto	Post-doc, PhD	Erasmus MC	Contributor to SAP
Milada C. Småstuen	Professor	Oslo Metropolitan University	Statistical advisor
Ørjan Nesse Vigdal	PhD-student	Oslo Metropolitan University	Contributor to SAP
Lise Kretz	PhD-student	Oslo Metropolitan University	Contributor to SAP

Signatures: Person writing SAP, senior statistician and chief investigator

_____	_____	_____
Rikke Munk Killingmo	**Milada Småstuen**	**Margreth Grotle**
Main author of the SAP	Senior statistician responsible	Principal Investigator

3. Study aim

The aim of this study is 1) to identify modifiable prognostic factors for high costs related to healthcare utilization among older people seeking primary care with a new episode of back pain and 2) to replicate the identified associations of modifiable prognostic factors in a similar cohort of older back pain patients.

4. Study design, population and method

Study design

This study will be carried out in two steps. First, modifiable prognostic factors will be identified in a prospective observational cohort study with one year of follow-up within a Norwegian primary care setting (the BACE-N). Next, a replication analysis of identified prognostic factors will be conducted in a prospective observational cohort study within a Dutch primary care setting (the BACE-D).

The BACE-N and the BACE-D studies are part of the international BACE consortium [1]. The BACE-N study (ClinicalTrials.gov Identifier: NCT04261309) was classified as a quality assessment study by the Norwegian Regional Committee for medical Research Ethics (reference no. 2014/1634/REK vest) and approved by the Norwegian Social Science Data Service (reference no. 42149) in 2015.

Study population and recruitment

BACE-N: Eligible patients are people 55 years of age or older who seek primary care (physiotherapist, chiropractor or GP) with a new episode of back pain (preceded by 6 months without visiting a primary care provider for similar complaints). Patients are excluded if they have difficulties completing the questionnaires (e.g. unable to speak, read or write in Norwegian) or if they have difficulties completing the physical examination (e.g. are wheelchair bound). Patients are recruited from physiotherapist, chiropractors and GPs working in Norwegian primary care between April 2015 and February 2020. Patients who meet the eligibility criteria and complete the consent to participate are included in the study.

BACE-D: Eligible patients were people over 55 years of age (n=675) who sought primary care (GPs) with a new episode of back pain (preceded by 6 months without visiting a primary care provider for similar complaints). Patients were excluded if they had difficulties completing the questionnaires (e.g. unable to speak, read or write in Dutch) or if they had difficulties completing the physical examination (e.g. are wheelchair bound). Patients were recruited from GPs working in Dutch primary care between March 2009 and September 2011. Patients who meet the eligibility criteria and complete the consent to participate were included in the study.

Method

BACE-N: At baseline all patients responded to a comprehensive questionnaire and went through a standardized physical examination conducted by local research assistants at test

stations established within each recruiting area. Follow-up questionnaires will be sent at 3, 6, and 12 months after inclusion for completion at home. All questionnaires are preferably completed electronically using the Infopad system, but paper versions will be available for patients not familiar with electronic data collection. All information will be stored and analysed securely through Service for sensitive data (TSD) at the University of Oslo, Norway.

BACE-D: At baseline all patients responded to a comprehensive questionnaire and went through a standardized physical examination. Follow-up questionnaires were sent (by e-mail or postal) at 3, 6, 9 and 12 months after inclusion.

Variables

Outcome variable

The outcome of this study is costs related to healthcare utilization aggregated for one year of follow up and dichotomized as high and low. Having high costs related to healthcare utilization is defined as patients with costs in the top 25th percentile [2, 3].

Healthcare utilization within the BACE-N and the BACE-D will be self-reported and include; consultation to healthcare professionals (type and frequency), number of diagnostic examinations (type and frequency), number of days of hospitalization and/or institutionalisation (only included in the BACE-N), back operations and use of back medication (both prescription and over-the-counter, type and frequency). All variables, except back operations, will be reported with a 3 month recall period at 3, 6, and 12 months follow-up for the BACE-N, at 3, 6, 9, and 12 months follow-up for the BACE-D. Back operations will be reported with a 12 month recall period at 12 months follow-up.

Healthcare utilization during the one year of follow-up will be described as shown in table 3 (BACE-N and BACE-D). The total cost of healthcare utilization will be estimated based on information presented in table 3 and unit costs of healthcare resources collected from national pricelists in Norway and the Nederlands (see table 1).

Potential modifiable prognostic factors

Potential modifiable prognostic factors are factors expected to have the potential to be modified *through healthcare system encounters* and therefore classified as modifiable. Potential modifiable prognostic factors of high-costs related to healthcare utilization are based on previous literature and will be measured at baseline.

- Pain severity [2-7] measured by the NRS
- Disability [2-6, 8] measured by the RMDQ
- Health-related quality of life [6, 7] measured by the SF36 using the physical and mental summary score
- Emotional well-being [2, 3, 8-10] measured by the CES-D
- Kinesiophobia [3, 10] measured by the FABQ-PA
- Comorbidity [11] measured by the SCQ
- Radiating pain below the knee [3] measured by the question “did your back pain radiate to your legs last week? If yes, how far down did the pain radiate last week?” categorized into yes or no

- Expectations of recovery measured with a five-point scale

The Numeric Rating Scale (NRS) will be used to measure average pain severity last week [12]. The NRS, scored from 0 (no pain) to 10 (maximum pain), has been widely used to evaluate pain and has proven to be preferable when examining low back pain patients [13], also for Norwegian patients [14].

The Roland Morris Disability Questionnaire (RMDQ) [15] will be used to measure disability. The RMDQ is a widely used back-specific patient-reported measure of pain-related disability (0 = no disability, 24 = totally disabled). The Norwegian version has been validated and found to have good measurement properties when used among patients with low back pain [14, 16].

The Short Form-36 Health Status Questionnaire (SF36) [17] will be used to assess health-related quality of life. The SF36 consist of 36 items. It measures health on eight multi-item dimensions, covering physical functioning, social functioning, role limitations (physical problems), role limitations (emotional problems), mental health, vitality, pain, and overall evaluation of health [17]. Data-completeness of the SF36 in the general population in Norway seems to strongly declined with increasing age [18]. Hence, caution should be exercised when assessing subjective health or employing the norms among subjects aged 70 years or over [18].

The Center for Epidemiologic Studies Depression Scale (CES-D) will be used to assess emotional well-being. The CES-D has been widely used in studies of late-life depression. Psychometric properties are generally favourable [19]. The Norwegian version of the CES-D has been used among older patients in order to measure depression symptoms [20].

The Fear Avoidance Beliefs Questionnaire, physical activity subscale (FABQ-PA) [21] will be used to assess kinesiophobia. The FABQ-PA consists of four questions aimed towards physical activity, scored on a 7-point ordinal scale, which are summed up to a sum score, ranging from 0 (no fear) to 24 (maximum fear). The questionnaire has been translated into Norwegian and has shown acceptable psychometric properties in Norwegian patients with low back pain [22].

The Self-Administered Comorbidity Questionnaire (SCQ) [23] will be used to assess comorbidity. The SCQ is a 14-item measure of comorbidity for clinical and health services research settings. An individual can receive a maximum of 3 points for each medical condition: 1 point for the presence of the problem, another point if he/she receives treatment for it, and an additional point if the problem causes a limitation in functioning. Because there are 12 defined medical problems and 3 optional conditions, the maximum score totals 45 points if the open-ended items are used and 36 points if only the close-ended items are used.

Potential covariates

Potential covariates will be included in the analyses based on previous literature and will be measured at baseline.

- Sex [4-6, 24, 25]

- Age [4, 6, 24, 25]
- Education level [8, 26] measured as the highest education completed, categorised into high vs low (low consists of up to high school and occupational high school)
- Employment status measured by the question “do you have a paying job?” categorized into yes or no
- Pain duration [2] measured by the question “how many days have you had your current back pain?”
- Pain history [5] measured by the question “have you had back pain before?” categorized into yes or no
- First healthcare provider [27]
- Costs related to healthcare utilization prior to inclusion

Healthcare utilization prior to inclusion will be measured (with variables as described above) at baseline, in the period from baseline to 6 and 12 weeks retrospectively, for the BACE-N and the BACE-D study respectively. The total cost of healthcare utilization will be estimated as described above.

Other variables

Included patients (BACE-N and BACE-D) will be described with respect to the following baseline characteristics: ethnicity, pain location and healthcare utilization prior to inclusion (see table 2). In addition, we have included the following potential prognostic factors and covariates (as described above): age, gender, educational level, employment status, first healthcare provider, pain severity, pain duration, pain history, disability, health-related quality of life, emotional well-being, kinesiophobia, expectation of recovery and comorbidity.

5. Statistical analyses

General analysis considerations

All analyses described in this plan 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 by a PhD-student using SPSS version 26 and controlled by a senior researcher/statistician. All statistical tests will be two-sided, and nominal p-values will be reported. 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. Missing data will be handled by multiple imputation, using 5 imputations and 10 iterations unless the missingness exceeds 30% and missing at random cannot be assumed. Fully conditioned specification method and regression estimation will be used. For variables where we are unable to use regression estimation due to computational difficulties, predictive mean matching will be used [28].

Description of study flow

The flow of participants through the study will be reported according to the STROBE guidelines [29] with a flow chart (see figure 1). Reasons for dropout will be provided where known. Differences between responders and non-responders will be evaluated.

Participant characteristics

Baseline characteristics of included patients will be presented as shown in table 2.

Preparatory analysis

First, type and frequency of use of different healthcare resources will be calculated for each of the follow-up periods; from baseline to 3 months, 3 to 6 months, and 9 to 12 months for the BACE-N study, and from baseline to 3 months, 3 to 6 months, 6 to 9 months, and 9 to 12 months for the BACE-D. Healthcare utilization will be presented as shown in table 3.

Next, costs will be estimated based on information presented in table 3, and unit costs of healthcare resources collected from national pricelists in Norway and the Nederlands (see table 1). Costs related to back medication will be estimated based on medication type (not exact medication name) and frequency of use. Data on dosage is not available. All costs will be presented in Euros (€) 2020. Costs of healthcare utilization will be described with median and interquartile range for the entire follow-up period as shown in table 4.

Identification analysis

Univariable and multivariable binary logistic regression models will be used to investigate individual association (crude and adjusted for selected covariates) between each predefined prognostic factor and costs related to healthcare utilization (within the BACE-N). 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). The results will be presented as crude and adjusted odds ratios (OR) with 95% confidence intervals (CI) as shown in table 5.

Replication analysis

Univariable and multivariable binary logistic regression models will be used, as described above, to replicate findings from the identification analysis within the BACE-D material. The results will be presented as crude and adjusted odds ratios (OR) with 95% confidence intervals (CI) as shown in table 5. The decision on whether findings 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 [30].

Sample size

This study contains secondary analyses embedded in the BACE-N and the BACE-D study. Details on sample size calculation are provided in the BACE-N (ClinicalTrials.gov Identifier: NCT04261309) and the BACE-D protocol [1].

To determine statistical power of this study we used number of events per variable (EPV) [46-50] and the rule-of-thumb of “10 events per 1 analysed variable” [51-54]. With a sample size of 450 participants within the BACE-N study, we anticipate 112 participants to be in the top 25th percentile of costs due to healthcare utilization and categorised as having high costs (yes/no) due to healthcare utilization (events). An EPV of 10 will allow a maximum of 11 prognostic variables to be included in

the final multivariable prediction model. With a sample size of 675 participants in the BACE-D, we anticipate 168 participants to be in the top 25th percentile of costs due to healthcare utilization and defined as having high costs (yes/no) due to healthcare utilization (events). An EPV of 10 will allow a maximum of 16 prognostic variables to be included in the final multiple prediction model.

Sensitivity analysis

To assess the robustness of the results complete case analysis (without using imputation for missing data) will be carried out as a sensitivity analysis.

6. Selection bias, information bias and covariates

Selection bias:

Because of limited resources and practical reasons related to recruitment from a broad network of clinicians, the BACE-N and the BACE-D lacks information on eligible study participants that declined to participate or for other reasons were not invited. In order to assess representativeness, the BACE-N study sample will be compared on key sociodemographic variables with a sample from the longitudinal population study of people in the second half of life; The Norwegian study on life course, ageing and generation (NORLAG). The NORLAG study is expected to represent a representative sample of older people with musculoskeletal complaints.

Response rate at each assessment point and reasons for loss to follow-up will be reported. Key baseline characteristics will be compared between those lost to follow-up and those remaining in the study.

Information bias:

To reduce the risk of information bias, the study outcome (costs) will be measured in an identical manner in all included cases, and in the best possible way within the framework of the BACE-N and the BACE-D study.

Covariates:

Covariates may influence associations between prognostic factors and outcome. Therefore, in line with the PROGRESS framework and recommendations for type 2 studies [31], we will adjust for covariates when evaluating prognostic factors.

Table 1 Cost categories, units, unit price, all numbers in Euros (€) for 2020?

Cost categories	Unit	Norwegian unit price (€)	Dutch unit price (€)	Reference (source)
Primary care				
General practitioner	Per visit			
Medical specialist	Per visit			
Occupational physician	Per visit			
Physiotherapist	Per visit			
Chiropractor	Per visit			
Manuel therapist	Per visit			
Naprapath	Per visit			
Osteopath	Per visit			
Psychologist	Per visit			
Other therapists	Per visit			
Back medication				
Paracetamol	Per daily defined dose			
NSAID	Per daily defined dose			
Muscle relaxant	Per daily defined dose			
Sleep medication	Per daily defined dose			
Cortisone	Per daily defined dose			
Opioid	Per daily defined dose			
Antidepressant	Per daily defined dose			
Anticonvulsant	Per daily defined dose			
Others	Per daily defined dose			
Examinations				
Blood sample	Per examination			
X-ray	Per examination			
MRI	Per examination			
CT	Per examination			
Others??	Per examination			
Secondary care				
Back operation	Per operation			
Hospitalization (non-operation)	Per day			
Rehabilitation stay	Per day			

NoMA, Norwegian Medicines Agency

Table 2 Patient characteristics and clinical status at baseline

	BACE-N (n = x)	BACE-D (n = x)
Female, N (%)		
Age in years, mean (SD)		
Education level high, N (%)		
Ethnicity Norwegian (BACE-N) or Dutch (BACE-D), N (%)		
Employment status, N (%)		
Currently paid work		
First healthcare provider, N (%)		
General practitioner		
Physiotherapist		
Chiropractor		
Pain location, N (%)		
Lumbar		
Thoracic		
Radiating pain below the knee		
Average pain severity last week (NRS 0-10), median (IQR)		
Pain duration, N (%)		
< 6 weeks		
6 weeks to 3 months		
> 3 months		
Previous episodes of back pain, N (%)		
Disability (RMDQ 0-24), mean (SD)		
Comorbidity (SCQ, 0-15)		
Health-related QOL (SF36, 0-100), mean (SD)		
Physical component		
Mental component		
Emotional well-being (CES-D 0-60)		
Kinesiophobia (FABQ-PA 0-24)		
Expectation of recovery within 3 months, N (%)		
Fully recovered		
Much better		
No change or worse		
Healthcare utilization prior to inclusion		
Patients with primary care consultation last 6 (BACE-N) or 12 (BACE-D) weeks, N (%)		
General practitioner		
Medical specialist		
Occupational physician		
Physiotherapist		
Chiropractor		
Manual therapist		
Naprapath		
Psychologist		
Other therapists		
Patients with use of back medication, N (%)?		
Patients with diagnostic examination last 6 (BACE-N) or 3 (BACE-D) months, N (%)		
Blood sample		
X-ray		
MRI/CT scan		
Patients with previous hospitalization, N (%)		
Patients with previous rehabilitation stay, N (%)		

NRS indicates Numeric Rating Scale; RMDQ, The Roland Morris Disability Questionnaire; SCQ, The Self-Administered Comorbidity Questionnaire; SF-36, 36-Item Short-Form Health Survey; CES-D, The Center for Epidemiologic Studies Depression Scale; FABQ-PA, The Fear Avoidance Beliefs Questionnaire, physical activity subscale.

Table 3 Healthcare utilization throughout one-year of follow-up

	BACE-N			BACE-D			
	0-3 months	>3-6 months	>9-12 months	0-3 months	>3-6 months	>6-9 months	>9-12 months
Primary care							
Patients with primary care consultation, N (%)							
General practitioner							
Medical specialist							
Occupational physician							
Physiotherapist							
Chiropractor							
Manual therapist							
Naprapath							
Psychologist							
Other therapists							
No. of general practitioner consultations, median (IQR)							
No. of medical specialist consultations, median (IQR)							
No. of occupational physician consultations, median (IQR)							
No. of physiotherapist consultations, median (IQR)							
No. of chiropractor consultations, median (IQR)							
No. of manual therapist consultations, median (IQR)							
No. of naprapath consultations, median (IQR)							
No. of psychologist consultations, median (IQR)							
No. of other consultations, median (IQR)							
Back medication							
Patients with use of back medication, N (%)							
Paracetamol							
NSAID							
Muscle relaxants							
Sleep medication							
Cortisone							
Opioid							
Others							
Frequency of use paracetamol, N (%)							
Daily							
Weekly							
Monthly or less							
Frequency of use NSAID, cortisone N (%)							
Daily							
Weekly							
Monthly or less							
Frequency of use muscle relaxants, sleep medication, N (%)							
Daily							
Weekly							
Monthly or less							
Examinations							
Patients with additional diagnostic examination, N (%)							
Blood sample							
X-ray							
MRI/CT scan							
Others							
Secondary care							
Patients with back operation, N (%)							
Patients with hospitalization, N (%)							
Duration of stay in days, median (IQR)							
Patients with rehabilitation stay, N (%)							
Duration of stay in days, median (IQR)							

Table 4 Cost related to healthcare utilization from 0-12 month*

	BACE-N	BACE-D
Primary care		
General practitioner		
Medical specialist		
Occupational physician		
Physiotherapist		
Chiropractor		
Manual therapist		
Naprapath		
Psychologist		
Other therapists		
Back medication		
Paracetamol		
NSAID, cortisone		
Muscle relaxants, sleep medication		
Examinations		
Blood sample		
X-ray		
MRI		
CT		
Others		
Secondary care		
Back operation		
Hospitalization and/or rehabilitation stay		
Total costs		

Values are median (interquartile range) of costs (€). *Cost related to healthcare utilization for the entire follow-up period is calculated on basis of the three (BACE-N) and four (BACE-D) follow-up periods

Table 5 Binary logistic regression analyses; individual associations between modifiable prognostic factors and high costs related to healthcare utilization (dependent variable)

	BACE-N		BACE-D	
	Crude OR (95% CI)	Adjusted OR* (95% CI)	Crude OR (95% CI)	Adjusted OR* (95% CI)
Pain severity (NRS, 0-10)				
Disability (RMDQ, 0-24)				
Health-related QOL (SF36, 0-100)				
Physical component				
Mental component				
Emotional well-being (CES-D 0-60)				
Kinesiophobia (FABQ-PA 0-24)				
Comorbidity (SCQ, 0-15)				
Radiating pain below the knee				
Yes				
No				
Expectation of recovery within 3 months, N (%)				

OR indicates odds ratio; CI, confidence interval; NRS, Numeric Rating Scale; RMDQ, The Roland Morris Disability Questionnaire; SF-36, 36-Item Short-Form Health Survey; CES-D, The Center for Epidemiologic Studies Depression Scale; FABQ-PA, The Fear Avoidance Beliefs Questionnaire, physical activity subscale; SCQ, The Self-Administered Comorbidity Questionnaire. *Adjusted by gender, age, education level, employment status, pain duration, pain history and costs related to healthcare utilization prior to inclusion.

SAP Modifiable prognostic factors of high cost related to healthcare utilization among older people seeking primary care with a new episode of back pain - an identification and replication study (working title)

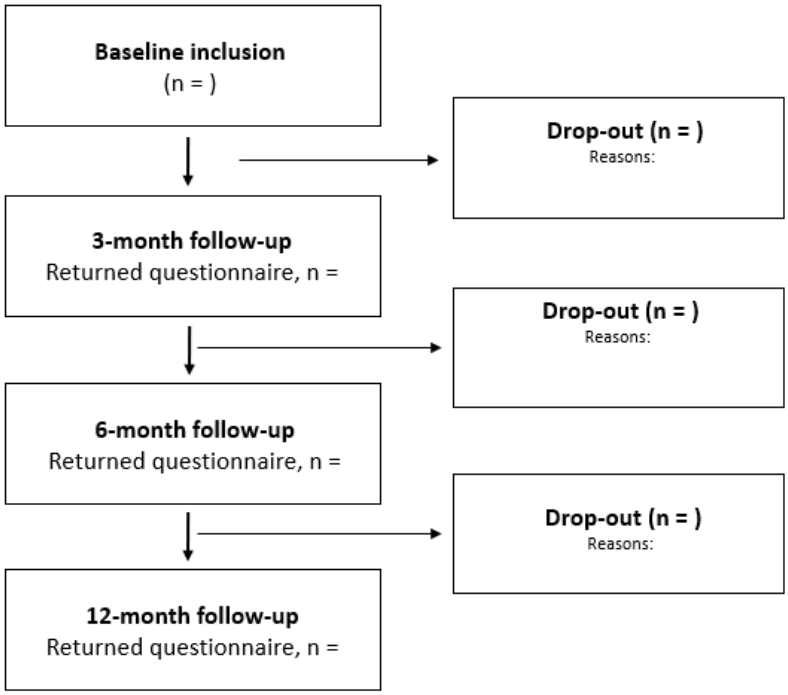


Figure 1. Flow chart of the study

References

1. Scheele, J., et al., *Back complaints in the elders (BACE); design of cohort studies in primary care: an international consortium.* BMC Musculoskelet Disord, 2011. **12**: p. 193.
2. 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.
3. 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.
4. Wenig, C.M., et al., *Costs of back pain in Germany.* Eur J Pain, 2009. **13**(3): p. 280-6.
5. Ferreira, M., et al., *Factors defining care-seeking in low back pain - A meta-analysis of population based surveys.* European Journal of Pain, 2010. **14**(7): p. 747.e1-747.e7.
6. Mutubuki, E.N., et al., *Predictive factors of high societal costs among chronic low back pain patients.* Eur J Pain, 2019.
7. Lentz, T.A., et al., *Factors associated with persistently high-cost health care utilization for musculoskeletal pain.* PLoS One, 2019. **14**(11): p. e0225125.
8. Lim, K.L., P. Jacobs, and S. Klarenbach, *A population-based analysis of healthcare utilization of persons with back disorders: results from the Canadian Community Health Survey 2000-2001.* Spine (Phila Pa 1976), 2006. **31**(2): p. 212-8.
9. Stewart, W.F., et al., *Patterns of health care utilization for low back pain.* J Pain Res, 2015. **8**: p. 523-35.
10. Keeley, P., et al., *Psychosocial predictors of health-related quality of life and health service utilisation in people with chronic low back pain.* Pain, 2008. **135**(1-2): p. 142-50.
11. Ritzwoller, D.P., et al., *The association of comorbidities, utilization and costs for patients identified with low back pain.* BMC Musculoskelet Disord, 2006. **7**: p. 72.
12. 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.
13. Strong, J., R. Ashton, and D. Chant, *Pain intensity measurement in chronic low back pain.* Clin J Pain, 1991. **7**(3): p. 209-18.
14. Grotle, M., J.I. Brox, and N.K. Vollestad, *Concurrent comparison of responsiveness in pain and functional status measurements used for patients with low back pain.* Spine (Phila Pa 1976), 2004. **29**(21): p. E492-501.
15. Roland, M. and R. Morris, *A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain.* Spine (Phila Pa 1976), 1983. **8**(2): p. 141-4.
16. Grotle, M., J.I. Brox, and N.K. Vollestad, *Cross-cultural adaptation of the Norwegian versions of the Roland-Morris Disability Questionnaire and the Oswestry Disability Index.* J Rehabil Med, 2003. **35**(5): p. 241-7.
17. Ware, J.E., Jr. and C.D. Sherbourne, *The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection.* Med Care, 1992. **30**(6): p. 473-83.
18. Loge, J.H. and S. Kaasa, *Short form 36 (SF-36) health survey: normative data from the general Norwegian population.* Scand J Soc Med, 1998. **26**(4): p. 250-8.
19. Radloff, L.S., *The CES-D scale: a Self-Report Depression Scale for Research in the General Population.* Applied psychological measurement, 1977. (1(3)): p. 385-355.
20. SE., T.K.C., *Funksjonshemning, ensomhet og depresjon: Hva betyr ensomhet for om personer med funksjonshemning opplever depresjon?* Tidsskrift Norsk Psykologforening, 2008. **1**.
21. Waddell, G., et al., *A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability.* Pain, 1993. **52**(2): p. 157-68.
22. Grotle, M., J.I. Brox, and N.K. Vollestad, *Reliability, validity and responsiveness of the fear-avoidance beliefs questionnaire: methodological aspects of the Norwegian version.* J Rehabil Med, 2006. **38**(6): p. 346-53.

23. Sangha, O., et al., *The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research.* Arthritis Rheum, 2003. **49**(2): p. 156-63.
24. Rattay, P., et al., *[Utilization of outpatient and inpatient health services in Germany: results of the German Health Interview and Examination Survey for Adults (DEGS1)].* Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz, 2013. **56**(5-6): p. 832-44.
25. Chechulin, Y., et al., *Predicting patients with high risk of becoming high-cost healthcare users in Ontario (Canada).* Healthc Policy, 2014. **9**(3): p. 68-79.
26. Hoebel, J., et al., *Socioeconomic Status and Use of Outpatient Medical Care: The Case of Germany.* PLoS One, 2016. **11**(5): p. e0155982.
27. Fritz, J.M., J. Kim, and J. Dorius, *Importance of the type of provider seen to begin health care for a new episode low back pain: associations with future utilization and costs.* J Eval Clin Pract, 2016. **22**(2): p. 247-52.
28. Van Buuren, S., *Flexible imputation of missing data*. 2018: CRC press.
29. von Elm, E., et al., *The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies.* PLoS Med, 2007. **4**(10): p. e296.
30. Kent, P., et al., *A conceptual framework for prognostic research.* BMC Med Res Methodol, 2020. **20**(1): p. 172.
31. Riley, R.D., et al., *Prognosis Research Strategy (PROGRESS) 2: prognostic factor research.* PLoS Med, 2013. **10**(2): p. e1001380.

Statistical analysis plan (SAP) for:

One-year clinical course of back-related disability and prognostic value of comorbidity on disability during one year follow-up in older patients visiting primary care with a new episode of back pain

Project:

BACk pain in Elders in Norway (BACE-N): a prospective cohort study of older people visiting primary care with a new episode of back pain

NCT identifier: NCT04261309
Document date: 22.03.2021

Table of Contents

Administrative information

Scope

This document is a supplement to the BACE-N protocol (ClinicalTrials.gov Identifier: NCT04261309) The current Statistical Analysis Plan has been written while data collection was ongoing (we had access to baseline data, but not to follow-up data) and it will be uploaded to the ClinicalTrials.gov before full access to the study database.

Working title

One-year clinical course of back-related disability and prognostic value of comorbidity on disability during one year follow-up in older patients visiting primary care with a new episode of back pain

Version of SAP

1.0

Ethical approval

The BACE-N study was deemed a "quality control project" by the Regional Ethical committee, and treatment by the ethical committee was thus not considered necessary as per 11.11.2014 (reference number: 2014/1634/REK vest)

Approval from the Norwegian Social Science Data Service was obtained on 02.03.2015 (reference number: 42149).

Names, affiliations and roles of SAP contributors

Name	**Title**	**Affiliation**	**Role**
Margreth Grotle	PT, professor, PhD	Oslo Metropolitan University P44, 0167 Oslo Phone: 90111172 Email: mgrotle@oslomet.no	Principal Investigator
Ørjan Nesse Vigdal	PT, PhD-student	Oslo Metropolitan University	Main author of SAP
Rikke Munk Killingmo	PT, PhD-student	Oslo Metropolitan University	Contributor to SAP
Kjersti Storheim	PT, professor, PhD	Oslo Metropolitan University	Contributor to SAP
Milada C. Småstuen	Professor, PhD	Oslo Metropolitan University	Statistical advisor
Lise Kretz	Chiro, PhD-student	Oslo Metropolitan University	Contributor to SAP

Study sponsor

The BACE-N-study has received funding from Oslo Metropolitan University, The Norwegian Fund for Post-Graduate Training in Physiotherapy and “Et liv i bevegelse” (A life in movement) – Norwegian chiropractors’ research foundation.

Signatures

_____	_____	_____
Ørjan Nesse Vigdal	Milada C. Småstuen	Margreth Grotle
Main author of SAP	Statistical advisor	Principal investigator

Introduction

Background and rationale for study

Back pain is common in all age groups (1), and one systematic review highlights that disabling back pain is more prevalent in older people than in younger people (2). The clinical course of back-related disability in older adults with back pain has not been extensively studied. Two studies suggests that improvements in disability were modest the first three months, with little to no improvements on group level after three months (3, 4). It is well documented that number of comorbidities are associated with the clinical course of back-related disability in older adults, but the prognostic value of comorbidity is still highly uncertain (5-11).

Study aim:

The primary aim of this study is to examine the clinical course of back-related disability measured at baseline, 3, 6 and 12 months after a new episode of back pain. The secondary aim is to assess the prognostic value of number and severity of comorbidity at baseline for changes in back-related disability over one year of follow-up.

Study design and method

Study design

BACE-N is based on the previously published BACE study protocol from the BACE international consortium (12). The BACE-N protocol has been published (ClinicalTrials.gov Identifier: NCT04261309).

The BACE-N study is a prospective observational cohort study with a 2-year follow-up time. This study will use data from baseline, 3, 6 and 12 months of follow-up. Design of the study was made within the framework PROGnosis RESearch Strategy (PROGRESS), which is a framework for ensuring and enhancing the quality of prognostic studies (13). The primary aim is relating overall prognosis, and the secondary aim is relating confirmatory prognostic factor research (13, 14).

Study population

Patients aged ≥55 years visiting a general practitioner (GP), physiotherapist or chiropractor for a new episode of back pain are invited to participate in the study. Back pain is defined as pain located in the region from the top of the scapula to the sacrum, with or without radiating leg pain. An episode of back pain is defined as “new” if the patient has not received health care for the same back complaint during the last 6 months. The exclusion criteria were: Difficulty completing the study questionnaires due to language or cognitive difficulties, mobility impairments impeding the clinical examination (wheelchair-bound patients), had received healthcare for the same back complaint during the last 6 months (notwithstanding care initiated within the previous 4 weeks from time of baseline assessment).

Data collection

At baseline, patients receive a clinical examination and questionnaire. Patients are then given a follow-up questionnaire either through email or mail at 3-months, 6-months and 12-months follow-up.

Description of treatment received during follow-up

The study participants continue their health care in agreement with their healthcare provider regardless of inclusion in the study. This means that patients may receive education and advice, exercise therapy, massage, manipulations, mobilizations, pharmacological therapy, or additional diagnostic testing and referrals, all of which constitutes "usual" primary care (15). Treatment given is at the discretion of the healthcare provider and the patient.

Variables

Outcome measure:

Back-related disability, measured using the Norwegian, validated version of the Roland-Morris Disability Questionnaire (RMDQ) (16, 17). This is a questionnaire with 24 statements regarding abilities to perform ADL tasks, with a dichotomous yes/no answer. The answers are summed to a total score ranging from 0-24, where 0 indicates no disability and 24 indicates "maximum" disability. RMDQ has been found to measure several dimensions of back-related disability (18). We plan to use RMDQ as a continuous measure, as recommended by the PROGRESS framework (14). This will ensure easier comparability to similar studies, and easier inclusion in future meta-analyses.

RMDQ is measured at baseline, 3 months, 6 months and 12 months.

Prognostic factor measurement:

The prognostic factor of interest is comorbidity measured at baseline with a modified version of the Self-administered Comorbidity Questionnaire (SCQ) (19). The original questionnaire measures 13 pre-specified comorbidities, and 2 non-specified. The item "back pain" has been removed for this study and replaced with an additional non-specified item. Thus, the count of comorbidity ranges from 0-15. The diseases listed are: Heart disease, high blood pressure, lung disease, diabetes, ulcer or stomach disease, kidney disease, liver disease, anemia or other blood disease, cancer, depression, osteoarthritis, rheumatoid arthritis and up to 3 non-specified comorbidities. The SCQ measures comorbidities on three levels: 1) Do you have the problem? 2) Do you receive treatment for it? 3) Does it limit your activities? All levels are answered on a dichotomous yes/no level, and you only answer level 2 and 3 if you have answered "yes" on level 1. An individual can receive a maximum of 3 points for each medical condition: 1 point for the presence of a comorbidity, 1 point if they receive treatment for the condition, and 1 point if the condition limits their functioning. The maximum score for the full SCQ is thus 45 points.

We plan to use SCQ part 1 comorbidity count (0-15 scale), and the full SCQ (0-45) in separate models. Previous studies have found linear relationships between comorbidity and back-related disability (7, 11), and thus we plan to treat these variables as linear. Linearity with outcome will be assessed, and deviations will be handled appropriately.

Covariates:
Covariates are presented in table 1. They are chosen based on being "established" prognostic factors in the literature, having been utilized in similar studies previously (to enhance comparability), and for being readily available and easy to measure in clinical practice.

Table 1: Covariates, measurement level and/or instrument, rationale for inclusion

Factor	Measurement level	Rationale
Age	Continuous, minimum age ≥55	Standard covariate in previous studies (7, 8, 10), and in some studies reported to be associated with disability levels in older adults (4, 6, 20)
Sex	Dichotomous	Standard covariate in previous studies (7, 8, 10), and in some studies reported to be associated with disability levels in older adults (6, 20)
BMI	Measured continuous, divided into categories: <20, 20-25, 25-30, <35	Covariate in previous study (7), and found to be associated with disability in older adults (4, 21)
Back pain duration	Measured in days, but categorized to an ordinal scale of 3 categories: 0-6 weeks, 6-12 weeks, >12 weeks, similar to other BACE studies (22)	Covariate in previous study (8), and found to be associated with disability in older adults (4, 6, 11, 20)
Baseline disability	Continuous. Roland-Morris disability questionnaire, 0-24 scale.	Standard covariate, given the outcome is disability. Only applicable if not using mixed models.
Pain severity last week	Continuous. Numeric rating scale 0-10.	Associated with disability in one study in older adults (4).
Expectation of recovery within the next three months	5-point ordinal Likert scale. From "I am fully recovered" to "I am worse than ever".	Consistently associated with disability in older adults (4, 6, 11)

Statistical power consideration

The published protocol for the BACE-N study estimates that we need a total of 450 patients included in the study. Allowing a 15% dropout rate at 12 months, we will have approximately 380 participants available at 12 months.

We used the power estimation tool in Stata to determine level of statistical power. With a sample size of 450, and an estimated r^2 of 0.30 for the full model of comorbidity adjusted for covariates, we have over 90% power to detect an r^2-change of 0.017 or higher when adding comorbidity to the model.

Handling of missing data

Missing data will be handled with multiple imputation, using 5 imputations and 10 iterations unless the missing exceeds 30% and missing at random cannot be assumed. We will use the fully conditioned specification method, and regression estimation. For variables where we are unable to use regression estimation due to computational difficulties, predictive mean matching will be used (23).

Missing values for RMDQ will be handled by replacing missing items with the mean of the answered items for the individual, if less than 30% of the items are missing.

A mixed model, which we intend to use for the primary and secondary aim, does not require variable level imputation for the outcome measure. However, missing values on item level on

the RMDQ might still be an issue. To solve this, we will replace missing items for RMDQ with the mean of the answered items for the individual, if less than 7 (30%) of the 24 items are missing.

Statistical analyses

General statistical considerations

Analyses described in this statistical analysis plan are considered a priori analyses. Possible post-hoc exploratory analyses will be explicitly identified in the article. All analyses will be carried out by the first author using Stata version 16, under supervision from the principal investigator and the advisory statistician. The statistical tests will be 2-sided, and p-value will be reported. If the p-value is less than 0.05, the test is deemed statistically significant. 95% confidence intervals will be reported on point estimates.

Statistical analyses

Description of study flow and study sample:

The flow of participants will be reported with a flow chart. Reasons for exclusion and loss to follow-up will be provided where known. Descriptive data of the study sample will be reported using mean and standard deviation for normally distributed continuous variables, median and interquartile range for variables with skewed distrubution, and with frequency and proportions for categorical variables. Normal distribution will be examined visually using histogram and QQ-plot, and statistically with the Kolmogorov-Smirnov test. Baseline characteristics will be presented in a table. See proposed tables and figures below.

Descriptive statistics will be used to present the mean and standard deviation if RMDQ is normally distributed, or median and interquartile range if RMDQ is not normally distributed, for each time point: Baseline, 3 months, 6 months, 12 months. This will also be presented graphically, similarly to van der Gaag et al (3). This graphical presentation will also be performed stratified for number of comorbidities.

A person may be a responder at one time-point and a non-responder at another. Therefore, an analysis of responders versus non-responders will be performed for each time point, using bivariate analysis for baseline characteristics (chi square test, Individual Samples T-test, or Mann Whitney U-test). Results from these analyses will be presented in text, and the table available in supplementary material.

Model choice and model building strategy:

According to the STRATOS initiative task force, a complete prespecifying of all aspects in model building is unrealistic in observational studies (24). Thus, the following will provide a framework for analysis and model building in this study, not a detailed recipe. Decisions regarding final choice of models will be made by the first author, the advisory statistician and the principal investigator.

Primary aim, clinical course of back-related disability:
Mixed models for repeated measures will be used to account for statistical dependencies. RMDQ is the dependent variable, and time and first contact health provider will be entered as fixed factors. The exact handling of time as a continuous or categorical variable depends on the distribution. Previous studies have found that the clinical course of a back episode is not linear over time (3, 25, 26). It is therefore reasonable to believe that time will have to be treated as a categorical variable, or by introducing a quadratic term, depending on its distribution. Choice of covariance matrix is dependent on data structure. Interaction between first contact health provider and time will be analysed.

Secondary aim, prognostic value of number and severity of comorbidities on the course of back-related disability:
We will fit separate models for count of comorbidity (SCQ part 1, 0-15 scale) and count and severity of comorbidity (full SCQ, 0-45). The steps are outlined below:

1. The univariate association between SCQ and RMDQ over time will be assessed with a mixed model for repeated measures. An interaction term for SCQ*time will be tested, and kept if statistically significant. We will present the crude regression estimates from this analysis.

2. A mixed model with all the covariates from Table 1 and RMDQ over time will be fitted. From this model, r^2 will be presented as a measure of prognostic value.

3. SCQ will be added to the model from step 2. The adjusted regression estimate, and the r^2-change from step 2 to step 3 will be presented.

Previous studies have found a linear relationship between number and severity of comorbidity and disability (7, 27). Linearity between SCQ and RMDQ will be assessed with a scatter-plot. In case of non-linearity between SCQ and RMDQ comorbidity count and disability score, the following alternatives will be discussed (24, 28):

- Categorization of SCQ
- Transformation (logarithmic or cubic) of SCQ
- Fractional polynomials for SCQ
- Spline functions for SCQ

Sensitivity analyses

Complete case analyses will be performed to assess possible bias introduced by the multiple imputation procedures.

Proposed tables and figures

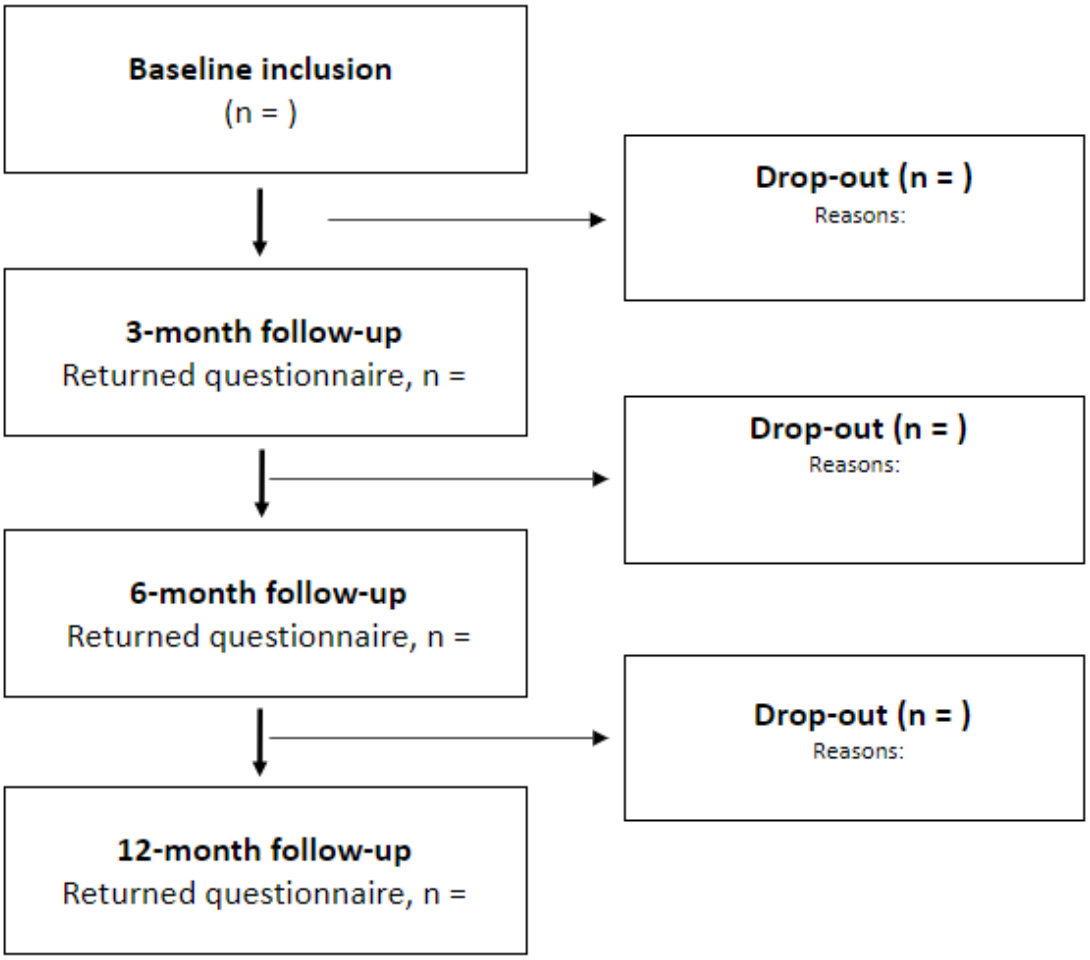


Figure: Flowchart of study participants

Baseline characteristics:

Table: Baseline characteristics

Variable	n	Values
Sociodemographic variables		
Age, median (IQR)		
Female, n (%)		
Married or living with partner, n (%)		
Education level high, n (%)		
General health variables		
Health-related quality of life (SF-36), median (range)		
Physical component score		
Mental component score		
Hazardous alcohol consumption (AUDIT-C), n (%)		
Smoking status, n (%)		
Current smoker		
Smoked previously		
Never		
Falls efficacy (FES-I), median (range)		
Back pain history and characteristics of current episode		
First healthcare provider		
General practitioner		
Physical therapist		
Chiropractor		
History of back pain, n (%)		
Using pain medication, n (%)		
Sleep problems weekly due to back pain, n (%)		
Back pain (NRS 0-10) (figure?), median (range)		
Disability (RMDQ 0-24) (figure?), median (range)		
Duration of current episode, n (%)		
0 – 6 weeks		
6 weeks – 3 months		
> 3 months		
Psychological variables		
Kinesiophobia (FABQ-PA 0-28), median (range)		
Depression (CES-D 0-60), median (range)		
Pain catastrophizing (PCS 0-52), median (range)		
Back beliefs and attitudes (BBQ 9-45), median (range)		
Expectations of back pain next 3 months (figure?), n (%)		
Fully recovered		
Much better		
No change or worse		
Psychosocial risk profile (SBT)		
Low risk		
Medium risk		
High risk		
Clinical variables		
Pain with active movements of the back, n (%)		
Positive radiculopathy diagnostic rule, n (%)		
Two or more red flags, n (%)		
Physical performance (BPS), median (range)		
Functional mobility (TUG), median (range)		

SF-36, Short Form health survey 36; AUDIT-C, alcohol use disorders identification test, score ≥3 for women and ≥4 for men ; FES-I, Falls Efficacy Scale – International; NRS, numeric rating scale; RMDQ, Roland-Morris Disability Questionnaire; FABQ-PA, Fear-Avoidance Beliefs Questionnaire – Physical Activity; CES-D, Center for Epidemiological Studies – Depression; PCS, Pain Catastrophizing Scale; BBQ, Back Beliefs Questionnaire; SBT, Start Back Tool; BPS, Back Performance Scale; TUG, Timed Up-and-Go.
*Comorbidities measured with Self-administered Comorbidity Questionnaire and a question on osteoporosis.

Appendix table: Baseline characteristics for responders vs non-responders for each time-point.

Variable	Responder 3 mo (n)	Non-responder 3 mo	Responder 6 mo (n)	Non-responder 6 mo (n)	Responder 12 mo (n=)	Non-responder 12 mo (n=)
Sociodemographic variables						
Age, median (IQR)						
Female, n (%)						
Married or living with partner, n (%)						
Education level high, n (%)						

General health variables	
Health-related quality of life (SF-36), median (range)	
Physical component score	
Mental component score	
Hazardous alcohol consumption (AUDIT-C), n (%)	
Smoking status, n (%)	
Current smoker	
Smoked previously	
Never	
Falls efficacy (FES-I), median (range)	
Back pain history and characteristics of current episode	
First healthcare provider	
General practitioner	
Physical therapist	
Chiropractor	
History of back pain, n (%)	
Using pain medication, n (%)	
Sleep problems weekly due to back pain, n (%)	
Back pain (NRS 0-10), median (range)	
Disability (RMDQ 0-24), median (range)	
Duration of current episode, n (%)	
0 – 6 weeks	
6 weeks – 3 months	
> 3 months	
Psychological variables	
Kinesiophobia (FABQ-PA 0-28), median (range)	
Depression (CES-D 0-60), median (range)	
Pain catastrophizing (PCS 0-52), median (range)	
Back beliefs and attitudes (BBQ 9-45), median (range)	
Expectations of back pain next 3 months (figure?), n (%)	
Fully recovered	
Much better	
No change or worse	
Psychosocial risk profile (SBT)	
Low risk	
Medium risk	
High risk	
Clinical variables	
Pain with active movements of the back, n (%)	
Positive radiculopathy diagnostic rule, n (%)	
Two or more red flags, n (%)	
Physical performance (BPS), median (range)	
Functional mobility (TUG), median (range)	

SF-36, Short Form health survey 36; AUDIT-C, alcohol use disorders identification test, score ≥3 for women and ≥4 for men ; FES-I, Falls Efficacy Scale – International; NRS, numeric rating scale; RMDQ, Roland-Morris Disability Questionnaire; FABQ-PA, Fear-Avoidance Beliefs Questionnaire – Physical Activity; CES-D, Center for Epidemiological Studies – Depression; PCS, Pain Catastrophizing Scale; BBQ, Back Beliefs Questionnaire; SBT, Start Back Tool; BPS, Back Performance Scale; TUG, Timed Up-and-Go.

*Comorbidities measured with Self-administered Comorbidity Questionnaire and a question on osteoporosis.

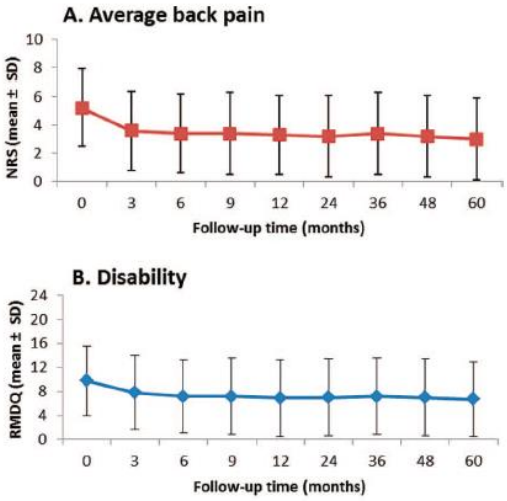
Prevalence of each comorbidity:

Table: Prevalence of each comorbidity at baseline

Comorbidity	**N**	**%**
Heart disease		
High blood pressure		
Lung disease		
Diabetes		
Ulcer or stomach disease		
Kidney disease		
Liver disease		
Anemia or other blood disease		
Cancer		
Depression		
Osteoarthritis		
Rheumatoid arthritis		
Osteoporosis*		
(Other non-prespecified comorbidities….)		

All comorbidities measured by Self-Administered Comorbidity Questionnaire

Example figure presenting clinical course, from van der Gaag et al (3):



Proportion of number of comorbidities among patients at baseline, example from Rundell et al (29):

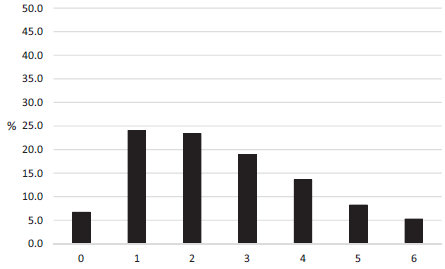


Figure 1. Proportion for number of additional pain sites among those with persistent back pain.

RMDQ middle value presented at each time point, presented based on number of comorbidities (29):

Table: Unadjusted RMDQ score at each timepoint in total cohort and by number of comorbidities

	Baseline	**3 months**	**6 months**	**12 months**
Total cohort	Xx	Xx	Xx	xx
No. of comorbidities				
0	Xx	Xx	Xx	xx
1				
2				
3				
4				
5				
6				

RMDQ = Roland-Morris Disability Questionnaire

Association between comorbidities and disability during 1-year follow-up:

Table: Estimates from **mixed model(XX)** of effect of comorbidities on RMDQ score during 1 year of follow-up

	β	**95% CI**	**R^2**
Comorbidities~			
Covariates			
Comorbidities w/covariates*			

RMDQ = Roland-Morris Disability Questionnaire
~unadjusted effect estimate
*adjusted for time, age, gender, education level, smoking status, hazardous alcohol consumption, back pain duration, probable radicular leg pain, (recruitment profession?)
^adjusted for age, gender, education level, smoking status, hazardous alcohol consumption, back pain duration, probable radicular leg pain, (recruitment profession?)

References

1. 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.
2. Dionne CE, Dunn KM, Croft PR. Does back pain prevalence really decrease with increasing age? A systematic review. Age and ageing. 2006;35(3):229-34.
3. van der Gaag WH, Enthoven WTM, Luijsterburg PAJ, van Rijckevorsel-Scheele J, Bierma-Zeinstra SMA, Bohnen AM, et al. Natural History of Back Pain in Older Adults over Five Years. The Journal of the American Board of Family Medicine. 2019;32(6):781-9.
4. Jarvik JG, Gold LS, Tan K, Friedly JL, Nedeljkovic SS, Comstock BA, et al. Long-term outcomes of a large, prospective observational cohort of older adults with back pain. Spine J. 2018;18(9):1540-51.
5. Scheele J, Enthoven WT, Bierma-Zeinstra SM, Peul WC, van Tulder MW, Bohnen AM, et al. Course and prognosis of older back pain patients in general practice: a prospective cohort study. Pain. 2013;154(6):951-7.
6. Deyo RA, Bryan M, Comstock BA, Turner JA, Heagerty P, Friedly J, et al. Trajectories of symptoms and function in older adults with low back disorders. Spine (Phila Pa 1976). 2015;40(17):1352-62.
7. Leopoldino AAO, Megale RZ, Diz JBM, Moreira BDS, Lustosa LP, Pereira LSM, et al. Influence of the number and severity of comorbidities in the course of acute non-specific low back pain in older adults: longitudinal results from the Back Complaints in the Elders (BACE-Brazil). Age and Ageing. 2020;49(1):96-101.
8. Rundell SD, Resnik L, Heagerty PJ, Kumar A, Jarvik JG. Comparing Performance of Comorbidity Indices in Predicting Functional Status, Health-Related Quality of Life, and Total Health Care Use in Older Adults With Back Pain. J Orthop Sports Phys Ther. 2019:1-18.
9. Rundell SD, Goode AP, Suri P, Heagerty PJ, Comstock BA, Friedly JL, et al. Effect of comorbid knee and hip osteoarthritis on longitudinal clinical and health care use outcomes in older adults with new visits for back pain. Archives of physical medicine and rehabilitation. 2017;98(1):43-50.
10. Rundell SD, Resnik L, Heagerty PJ, Kumar A, Jarvik JG. Performance of the Functional Comorbidity Index (FCI) in Prognostic Models for Risk Adjustment in Patients With Back Pain. PM&R. 2020.
11. van der Gaag WH, Chiarotto A, Heymans MW, Enthoven WTM, van Rijckevorsel-Scheele J, Bierma-Zeinstra SMA, et al. Developing clinical prediction models for non-recovery in older patients seeking care for back pain: the BACE prospective cohort study. PAIN. 2020;Articles in Press.
12. Scheele J, Luijsterburg PA, Ferreira ML, Maher CG, Pereira L, Peul WC, et al. Back complaints in the elders (BACE); design of cohort studies in primary care: an international consortium. BMC Musculoskelet Disord. 2011;12:193.
13. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. BMJ. 2013;346:e5595.
14. Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. PLoS Med. 2013;10(2):e1001380.
15. Kamper SJ, Logan G, Copsey B, Thompson J, Machado GC, Abdel-Shaheed C, et al. What is usual care for low back pain? A systematic review of health care provided to patients with low back pain in family practice and emergency departments. PAIN. 2020;161(4):694-702.
16. Roland M, Morris R. A Study of the Natural History of Back Pain: Part IDevelopment of a Reliable and Sensitive Measure of Disability in Low-Back Pain. spine. 1983;8(2):141-4.
17. Grotle M, Brox J, Vollestad N. Cross-cultural adaptation of the Norwegian versions of the Roland-Morris Disability Questionnaire and the Oswestry Disability Index. Journal of Rehabilitation Medicine. 2003;35(5):241-7.

18. Grotle M, Brox JI, Vøllestad NK. Functional Status and Disability Questionnaires: What Do They Assess?: A Systematic Review of Back-Specific Outcome Questionnaires. Spine. 2005;30(1):130-40.
19. Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN. The self-administered comorbidity questionnaire: A new method to assess comorbidity for clinical and health services research. Arthritis & Rheumatism. 2003;49(2):156-63.
20. Jarvik JG, Comstock BA, Heagerty PJ, Turner JA, Sullivan SD, Xu S, et al. Back pain in seniors: the back pain outcomes using longitudinal data (BOLD) cohort baseline data. BMC Musculoskeletal Disorders. 2014;15(1):1-23.
21. Jesus-Moraleida FR, Ferreira PH, Ferreira ML, Silva JP, Maher CG, Enthoven WT, et al. Back Complaints in the Elders in Brazil and the Netherlands: a cross-sectional comparison. Age and ageing. 2017;46(3):476-81.
22. Scheele J, Enthoven WT, Bierma-Zeinstra SM, Peul WC, van Tulder MW, Bohnen AM, et al. Characteristics of older patients with back pain in general practice: BACE cohort study. Eur J Pain. 2013;18(2):279-87.
23. Van Buuren S. Flexible imputation of missing data: CRC press; 2018.
24. Sauerbrei W, Perperoglou A, Schmid M, Abrahamowicz M, Becher H, Binder H, et al. State-of-the-art in selection of variables and functional forms in multivariable analysis--outstanding issues. arXiv preprint arXiv:190700786. 2019.
25. Itz CJ, Geurts JW, Kleef M, Nelemans P. Clinical course of non-specific low back pain: A systematic review of prospective cohort studies set in primary care. European Journal of Pain. 2013;17(1):5-15.
26. Costa LdCM, Maher CG, Hancock MJ, McAuley JH, Herbert RD, Costa LO. The prognosis of acute and persistent low-back pain: a meta-analysis. Canadian Medical Association Journal. 2012;184(11):E613-E24.
27. Bruusgaard D, Tschudi-Madsen H, Ihlebæk C, Kamaleri Y, Natvig B. Symptom load and functional status: results from the Ullensaker population study. BMC Public Health. 2012;12(1):1085.
28. Riley RD, Moons KG, Hayden JA, Sauerbrei W, Altman DG. Prognostic Factor Research. In: Riley RD, van der Windt DA, Croft P, Moons KG, editors. Prognosis Research in Healthcare Concepts, Methods and Impact. Oxford, United Kingdom: Oxford University Press; 2019.
29. Rundell SD, Patel KV, Krook MA, Heagerty PJ, Suri P, Friedly JL, et al. Multisite Pain Is Associated with Long-term Patient-Reported Outcomes in Older Adults with Persistent Back Pain. Pain Med. 2019.

Statistical analysis plan (SAP) for:

External validation of clinical prediction models for non-recovery in older patients seeking primary care for back pain

SAP Authors:

Ørjan Nesse Vigdal, Kjersti Storheim, Lise Kretz, Tarjei Rysstad, Are Hugo Pripp, Rikke Munk Killingmo, Wendelien van der Gaag, Alessandro Chiarotto, Bart Koes, Margreth Grotle (order not decided)

Project:

BACk pain in Elders in Norway (BACE-N): a prospective cohort study of older people visiting primary care with a new episode of back pain

NCT identifier: NCT04261309
Document date: 28.09.2022

Table of Contents

Administrative information

Scope

This document is a supplement to the BACE-N protocol (ClinicalTrials.gov Identifier: NCT04261309) The current SAP was written after data collection was completed, but registered in ClinicalTrials.gov before retrieving data from the database for analyses.

Working title

External validation of clinical prediction models for non-recovery in older patients seeking primary care for back pain

Version of SAP

1.0

Ethical approval

The BACE-N study was deemed a "quality control project" by the Norwegian Regional Committee for Medical Research Ethics, and treatment by the ethical committee was thus not considered necessary as per 11.11.2014. Reference: 2014/1634/REK vest

Approval from the Norwegian Social Science Data Service was obtained on 02.03.2015 (reg. no. 42149).

Names, affiliations and roles of SAP contributors (order not decided)

Name	**Title**	**Affiliation**	**Role**
Margreth Grotle	PT, professor, PhD	Oslo Metropolitan University P44, 0167 Oslo Phone: 90111172 Email: mgrotle@oslomet.no	Principal Investigator
Ørjan Nesse Vigdal	PT, PhD-student	Oslo Metropolitan University	Main author of SAP
Rikke Munk Killingmo	PT, PhD-student	Oslo Metropolitan University	Contributor to SAP
Kjersti Storheim	PT, professor, PhD	Oslo Metropolitan University	Contributor to SAP
Are Hugo Pripp	Statistician, PhD	Oslo Metropolitan University	Statistical advisor
Lise Kretz	DC, PhD-student	Oslo Metropolitan University	Contributor to SAP
Tarjei Rysstad	PT, PhD-student	Oslo Metropolitan University	Contributor to SAP

Wendelien van der Gaag	MD, PhD-student	Erasmus University Medical Center Rotterdam	Contributor to SAP
Bart Koes	Professor, PhD	Erasmus University Medical Center Rotterdam	Contributor to SAP
Alessandro Chiarotto	PT, PhD, post-doc	Erasmus University Medical Center Rotterdam	Contributor to SAP

Study sponsor

The BACE-N-study has received funding from Oslo Metropolitan University, The Norwegian Fund for Post-Graduate Training in Physiotherapy (grant number 90749) and “Et liv i bevegelse”, the Research Unit of the Norwegian Chiropractic Association.

Signatures

_____	_____	_____
Ørjan Nesse Vigdal	Are Hugo Pripp	Margreth Grotle
Main author of SAP	Statistical advisor	Principal investigator

Introduction

Background and rationale for study

Back pain remains the number one cause for years lived with disability globally (1). The prevalence of back pain is highest among adults between 40 and 80 years of age (2). This, in addition to an ageing population globally, has recently led to increased attention towards back pain in older adults (3, 4). Persistent back pain frequently have severe negative consequences for the older individual, health care systems and society (5). Thus, identifying those with higher risk of persistent levels of back pain or disability is of importance.

Prognostic models are a way of providing individual prognosis, using a standardized measurement tool (6). These models combine relevant prognostic factors to try to give an accurate prediction on the individual's prognosis (6). Most previous prognostic models or screening tools for back pain are developed for younger adults (7-11). Only one study has developed prognostic models for use in older adults with back pain (12). This study, performed in the Back Complaints in the Elders (BACE) – Netherlands study, found that a combination of biopsychosocial factors was able to accurately predict persistent back pain, persistent disability and persistent self-reported non-recovery (12). This model has yet to be externally validated, which is a crucial step in prognostic model research (13, 14). The BACE consortium, where research groups from different countries use similar methods and measurement tools to study back pain in older adults in primary care (3), offers an ideal setting for external validation of these prognostic models.

Study aim

The aim of this study is to externally validate three clinical prediction models for non-recovery from a new episode of back pain in older adults. If model performance is poor, the secondary aim is to recalibrate or update the models.

Here we present the study's Statistical Analysis Plan (SAP), which details the analysis steps for the external validation, to minimize the risk of data selection bias and data-driven interpretation of results.

Study design and method

Study design

BACE-N is based on the previously published BACE study protocol from the BACE international consortium (3). The BACE-N protocol has been published elsewhere (ClinicalTrials.gov Identifier: NCT04261309).

The BACE-N study is a prospective observational cohort study with a 2-year follow-up time. Patients are followed up after 3, 6, 12 and 24 months. This article will use data from baseline, 6 and 12 months of follow-up, in line with the original models developed in the BACE Netherlands study. The design of the current study was informed by the framework PROGnosis RESearch Strategy (PROGRESS), which is a framework for ensuring and enhancing the quality of prognostic studies (15). In particular, this is a PROGRESS type 3 study (6). Reporting will be according to the Transparent Reporting of a multivariable

prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement for prediction studies (16).

Study population

Development sample, BACE-D:
Patients aged ≥55 years visiting a general practitioner (GP) in the Netherlands with a new episode of back pain were invited to participate. Back pain was defined as pain located in the region from the top of the scapula to the sacrum, with or without radiating leg pain. An episode of back pain was defined as "new" if the patient has not received health care for the same back complaint during the last 6 months.

Exclusion criteria:

- difficulty completing the study questionnaires due to language or cognitive difficulties
- mobility impairments impeding the clinical examination (wheelchair-bound patients)
- had received healthcare for the same back complaint during the last 6 months.

Between March 2009 and September 2011, 675 patients were included.

Validation sample, BACE-N:
Patients aged ≥55 years visiting a GP, physiotherapist or chiropractor in Norway for a new episode of back pain are invited to participate in the study. The inclusion and exclusion criteria were identical to the development sample. Between April 2015 and February 2020, 452 patients were included in the study.

Data collection in validation sample

Participants were recruited from GP practices, physiotherapy practices and chiropractic practices. Healthcare providers were asked to invite consecutive patients, but given obvious time constraints in clinical practice they were not asked to collect data on patients that either declined to participate or for other reasons were not invited to participate in the study. Eligible patients received oral and written information during or after their consultation. Then the patients were either contacted by a study coordinator or contacted a study coordinator directly for assessment of eligibility. To facilitate the recruitment process, advertisements in media were used for inviting eligible patients to directly contact a study coordinator. These patients were screened for eligibility by the study coordinators (ØNV, RMK and LK) applying the same inclusion and exclusion criteria as mentioned above.

The study participants in BACE-N continue their health care in agreement with their healthcare practitioner regardless of inclusion in the study. After the initial clinical examination and questionnaire, patients were given a follow-up questionnaire either through email or mail at 3-months, 6-months, 12-months and 24-months follow-up. For this study we are using baseline data, and from 6 and 12-months follow-up.

Variables

Outcome measures:

Similarly to the development study (12), non-recovery is defined for three different outcomes:

- Persistent disability: A score of ≥4/24 on the Roland-Morris Disability Questionnaire (RMDQ) at both 6 and 12 months (17, 18). The RMDQ ranges from 0, no disability, to 24, maximum disability.
- Persistent back pain: A score of ≥3/10 on the Numeric Rating Scale (NRS) at both 6 and 12 months. The NRS ranges from 0, no pain, to 10, worst maximal pain.
- Persistent self-reported non-recovery: A score of ≥3 on a 7-point Global Perceived Effect (GPE) scale at both 6 and 12 months. GPE scores were:
    1. Full recovery
    2. Much better
    3. Somewhat better
    4. No change
    5. Somewhat worse
    6. Much worse
    7. Worse than ever

Prognostic factors:

In the development study there were 45 potential predictors. After an extensive literature review, consensus meetings, and sample size calculation, 14 candidate predictors were selected:

Table 1: Candidate prognostic factors from development study.

Prognostic factor	Measurement level	Measured with	Parameters in model
Age	Continuous		1
Sex	Binary		1
Chronic duration of current episode of back pain	Binary	Duration >3 months at baseline	1
Back pain intensity in the past week	Continuous, 0-10	NRS	1
Back-related disability	Continuous, 0-24	RMDQ	1
Recent episode of back pain	Binary	An episode of back pain within the past 6 months, for which they did not seek care	1
Musculoskeletal comorbidity	Binary	Modified Dutch version of SCQ	1
Radiating pain to the leg	Binary	Presence of radiating leg pain in either or both leg(s)	1
Spinal morning stiffness longer than 30 minutes	Binary	Modified KOOS-question 6a	1

Pain during spinal rotation	Binary	Pain with rotation of the upper body during clinical examination	1
Expectations of recovery	Binary.	Likert scale ranging from 1, “fully recovered” to 5, “worse than ever”. Dichotomized into “expecting improvement” and “not expecting improvement”.	1
Depressive symptomology	Continuous, 0-60	CES-D	1
Kinesiophobia	Continuous, 0-28	FABQ-PA	1
Pain catastrophizing	Continuous, 0-52	PCS	1
NRS; Numeric rating scale, RMDQ; Roland-Morris Disability Questionnaire, SCQ; Self-administered Comorbidity Questionnaire, KOOS; Knee injury and Osteoarthritis Outcome Score, CES-D; Center for Epidemiological Studies – Depression, FABQ-PA; Fear-Avoidance Beliefs Questionnaire – Physical Activity subscale, PCS; Pain Catastrophizing Scale.			

The published clinical prediction models:

The penalized models from the development samples are as follows:

Persistent back pain = -4.23 + 0.03*Age + 0.73*Chronic duration + 0.23*Back pain intensity – 0.10*Back pain intensity (cubic) spline + 0.07*Disability + 0.71*Recent episode + 0.38*Spinal morning stiffness + 0.42*Pain during spinal rotation – 0.79*Recovery expectation

Persistent disability = -5.97 + 0.04*Age + 0.63*Chronic duration + 0.32*Disability – 0.17*Disability (cubic) spline + 0.40*Recent episode + 0.52*Musculoskeletal comorbidity + 0.52*Spinal morning stiffness -0.83*Recovery expectation +0.08*Depressive symptomology - 0.08*Depressive symptomology (cubic) spline + 0.02*Pain catastrophizing

Self-reported non-recovery = -3.46 + 0.03*age + 0.60*Chronic duration + 0.13*Disability – 0.07*Disability (cubic) spline + 1.11*Recent episode + 0.36*Pain during spinal rotation - 0.95*Recovery expectation

Changes to candidate prognostic factors and spline terms from development to validation

The development study used a modified version of the Self-Administered Comorbidity Questionnaire (SCQ) which was unavailable at the time of designing the external validation study. Thus, the validation study uses a Norwegian version of the original SCQ.

Due to using different versions of the SCQ, we are unable to replicate the exact measurement of “musculoskeletal comorbidity” in the validation study. Thus, we will exchange this factor

with “comorbidity” measured with the SCQ as a binary variable with cut-off at ≥1 comorbidity.

Unfortunately, details on the cubic spline terms in the published original models were unavailable in the manuscript, and sufficient details from the development of the spline terms were unavailable to the authors of this SAP. This means that we are currently unable to replicate the spline terms from the development models in the external validation sample. Sufficient details on the spline terms could become available at a later time and may thus be incorporated before submission of the article based on these analyses.

Finally, there was a slight difference in wording for the prognostic factor “Spinal morning stiffness”. In the Netherlands, patients were asked to indicate if the stiffness lasted longer than 30 minutes, whereas in Norway, they were asked to indicate if it lasted longer than 60 minutes.

Sample size calculation

To ensure that the regression analyses are sufficiently powered for model extension, we used recent recommendations by Riley et al (19), and their online sample size calculator available at https://riskcalc.org/pmsamplesize/. Based on previous studies, we expect a prevalence of non-recovery for all outcomes of around 50% (12, 20). It has been suggested that 100-200 events are sufficient in validation studies (13). We assumed an outcome prevalence of 40%, a maximum of 14 candidate predictor parameters and C-statistic values of 0.8, similar to the optimism-adjusted C-statistic values from the development study. This yielded a minimum sample size of 426 participants.

Statistical analyses

General statistical considerations

Analyses described in this statistical analysis plan are considered a priori planned analyses. Possible post-hoc exploratory analyses will be explicitly identified in the article. All analyses will be carried out by the first author (ØNV) using Stata version 16 (StataCorp LLC, Texas, USA), under supervision from the principal investigator (MG) and the advisory statistician (AHP). The statistical tests will be 2-sided, and p-value will be reported. If the p-value is less than 0.05, the test is deemed statistically significant. 95% confidence intervals will be reported on effect sizes.

Handling of missing data

Development sample:

Multiple imputation with 25 imputations and 50 iterations, using predictive mean matching, was performed on predictors and outcomes, except for RMDQ, where missingness was accounted for by multiplying the total score of the individual by 100, and dividing by the number of missing items (12, 21). This 0-100 score was then re-fitted to the original 0-24 scale (divided the score by 4.167) of the RMDQ.

Validation sample:
Missing data on RMDQ at item-level at 6 and 12 months will be handled by replacing the missing item with the mean score for that individual, if less than 30% of the RMDQ items is missing. This yields identical results to the proposed procedure by Kent et al (21).

Multiple imputation may become necessary if there is substantial missing (>5%), which will yield insufficient power in the final models. The risk of this is present because the outcome measures used in the models require the participants to have valid answers at both 6 and 12 months. If multiple imputation of outcome becomes necessary, it will be performed on the outcome's original scale, before dichotomizing the outcomes and combining the time points. We will use all predictors and auxiliary variables in the multiple imputation models to reduce bias introduced in the models. Observed and imputed values will be compared in tables and plots, and convergence will be checked. Number of imputations needed is dependent on if convergence is reached, and percentage of participants with missing values (22, 23).

Analyses

Descriptive statistics:
Descriptive data will be reported for both the development and validation sample, using mean and standard deviation for normally distributed continuous variables, median and interquartile range for continuous variables with skewed distribution, and with frequency and proportions for categorical variables. Normal distribution will be examined visually using histogram and QQ-plot, and statistically with the Shapiro-Wilk test.

The proportion of patients with non-recovery will be presented for each outcome for the development and validation samples.

Loss-to-follow-up and dropout will be counted and assessed. Reasons for such will be presented where available. Patients withdrawing their informed consent during follow-up will not be included in the analyses. The baseline characteristics of patients lost to follow-up will be compared to those remaining in the study.

External validation:
We will follow the three steps outlined by Debray et al (24) when assessing external validity:

Step 1, sample relatedness:
Sample relatedness will be assessed by comparing the patient characteristics in the validation sample to the development sample. We will assess each characteristic separately, using summary measures like percentage, mean and standard deviation, or median and range. (Table 3 below.)

Step 2, external validation performance:
The models tested in the external validation sample looks as follows:

Persistent back pain = -4.23 + 0.03*Age + 0.73*Chronic duration + 0.23*Back pain intensity + 0.07*Disability + 0.71*Recent episode + 0.38*Spinal morning stiffness + 0.42*Pain during spinal rotation – 0.79*Recovery expectation

Persistent disability = -5.97 + 0.04*Age + 0.63*Chronic duration + 0.32*Disability + 0.40*Recent episode + 0.52*Comorbidity + 0.52*Spinal morning stiffness -0.83*Recovery expectation +0.08*Depressive symptomology + 0.02*Pain catastrophizing

Persistent self-reported non-recovery = -3.46 + 0.03*Age + 0.60*Chronic duration + 0.13*Disability + 1.11*Recent episode + 0.36*Pain during spinal rotation - 0.95*Recovery expectation

Model performance will be measured with calibration-in-the-large (CITL), calibration slope, Nagelkerke's pseudo-R^2, and Area under Receiver Operator Characteristics curve (AUC). (Table 4 and Figure 1 below.) CITL and calibration slope is assessed through a calibration plot, where observed outcomes are plotted against predicted risks. Perfect calibration is indicated by a CITL value of 0 and a calibration slope value of 1 (14). Nagelkerke's pseudo-R^2 is a measure of overall model performance that ranges between 0 (worst possible performance) and 1 (best possible performance). AUC is a measure of the model's discriminative performance. A value of 1 indicates perfect discrimination and a value of 0.5 indicates discrimination no better than chance. AUC values between 0.7 and 0.8 are considered acceptable, between 0.8 and 0.9 excellent, and between 0.9-1.0 outstanding (25).

Step 3, model recalibration and updating:
If the models perform unsatisfactorily in the validation sample, updating the models may be needed. We will assess several alternatives in a step-by-step manner during this stage, following the frameworks of Debray et al (24) and Steyerberg (14):

1. Adjustment of model intercept, in case of poor calibration-in-the-large.
2. Overall adjustment of the calibration slope, in case of poor calibration slope.
3. In case of poor discrimination in validation models compared to original models, model revision may be performed. This involves re-estimation of predictor effects, including assessing linearity assumptions of continuous prognostic factors with restricted cubic splines or fractional polynomials. The re-estimated predictor effects will be corrected for over-optimism by applying a uniform shrinkage factor derived from 500 bootstrapped samples.
4. Model extension, introducing new candidate predictors to the model and assess improvement in performance. The candidate predictor to be tested and rationale can be found in Table 2 below.

Table 2: Candidate prognostic factor for model extension.

Candidate prognostic factor	**Rationale**
Reduced sleep quality attributable to back pain (based on item 5i from PSQI)	Sleep problems are common in older adults (26), and associated with disability in a cross-sectional study on back pain in older adults (27). Reduced sleep quality often co-occurs with chronic pain (28-30).

PSQI: Pittsburgh Sleep Quality Index

If model performance after completing a step is acceptable, we will not proceed with further steps.

Sensitivity analyses

Complete case analyses will be performed to assess the possible bias introduced from the imputation procedure.

Proposed tables and figures

Baseline characteristics:

Table 3: Patient characteristics and clinical status at baseline in development sample and validation sample.

	Development sample, BACE-D		Validation sample, BACE-N	
	All participants (n=675)	**Missing, n (%)**	**All participants (n=452)**	**Missing, n (%)**
Age, mean (SD)				
Female, n (%)				
BMI, mean (SD)				
Education level high, n (%)				
Employment status currently paid work, n (%)				
Health-related quality of life (SF-36), mean (SD)				
Physical component score				
Mental component score				
Current smoker, n (%)				
Musculoskeletal comorbidity, n (%)				
One or more comorbidities, n (%)				
One or more previous back pain episodes, n (%)				
Recent episode of back pain, n (%)				
Radiating leg pain, n (%)				
Using pain medication, n (%)				
Sleep problems weekly due to back pain, n (%)				
Average back pain intensity last week (NRS, 0-10), mean (SD)				
Disability (RMDQ, 0-24), mean (SD)				
Duration of current episode > 3 months, n (%)				
Kinesiophobia (FABQ-PA, 0-24), mean (SD)				
Depression (CES-D, 0-60), mean (SD)				
Pain catastrophizing (PCS, 0-52), mean (SD)				
Expectations of being fully recovered or much better from back pain within next 3 months, n (%)				
Spinal morning stiffness more than 30/60 minutes[a], n (%)				
Pain with active rotation of the back, n (%)				

SF-36, Short Form health survey 36; AUDIT-C, alcohol use disorders identification test, score ≥3 for women and ≥4 for men; NRS, numeric rating scale; RMDQ, Roland-Morris Disability Questionnaire; FABQ-PA, Fear-Avoidance Beliefs Questionnaire – Physical Activity; CES-D, Center for Epidemiological Studies – Depression; PCS, Pain Catastrophizing Scale; BBQ, Back Beliefs Questionnaire; TUG, Timed Up-and-Go.

[a]Duration of morning stiffness was set at 30 minutes in BACE-D, and 60 minutes in BACE-N.

Table 4: Performance measures of external validation (This table will be provided for both the original models, and for updated/extended models)		
Persistent back pain (≥3/10 on NRS at both 6 and 12 months)		
Aspect	Measure	Value
Overall performance	Nagelkerke pseudo-R^2	
Discrimination	Area under the Curve	
Calibration	Calibration in-the-large	
	Calibration slope	
Persistent back-related disability (≥4/24 on RMDQ at both 6 and 12 months)		
Aspect	Measure	Value
Overall performance	Nagelkerke pseudo-R^2	
Discrimination	Area under the Curve	
Calibration	Calibration in-the-large	
	Calibration slope	
Persistent non-recovery (≥3/7 on GPE at both 6 and 12 months)		
Aspect	Measure	Value
Overall performance	Nagelkerke pseudo-R^2	
Discrimination	Area under the Curve	
Calibration	Calibration in-the-large	
	Calibration slope	

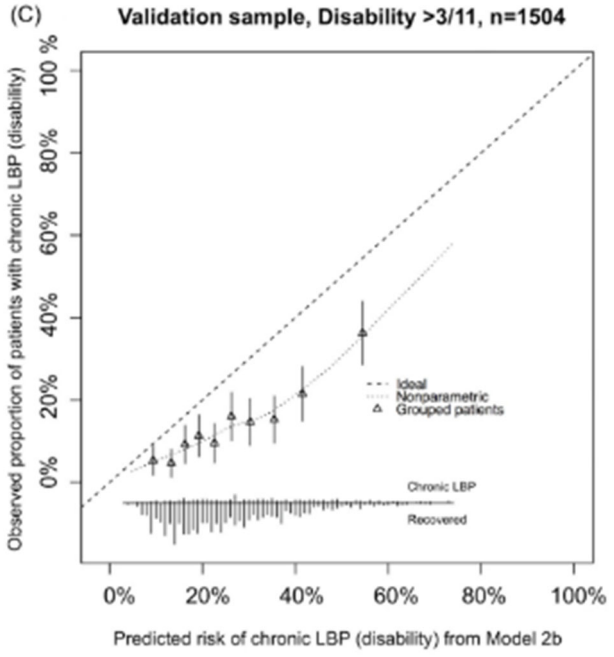


Figure 1: Calibration plot from Traeger et al (8). Calibration plots will be shown for the models tested in Step 2, and for possible updated and/or extended models (Step 3).

References

1. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet. 2012;380(9859):2163-96.
2. 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.
3. Scheele J, Luijsterburg PA, Ferreira ML, Maher CG, Pereira L, Peul WC, et al. Back complaints in the elders (BACE); design of cohort studies in primary care: an international consortium. BMC Musculoskelet Disord. 2011;12:193.
4. Jarvik JG, Comstock BA, Bresnahan BW, Nedeljkovic SS, Nerenz DR, Bauer Z, et al. Study protocol: the back pain outcomes using longitudinal data (BOLD) registry. BMC musculoskeletal disorders. 2012;13(1):64.
5. Ferreira ML, de Luca K. Spinal pain and its impact on older people. Best Practice & Research Clinical Rheumatology. 2017.
6. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10(2):e1001381.
7. Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, et al. A primary care back pain screening tool: Identifying patient subgroups for initial treatment. Arthritis & Rheumatism. 2008;59(5):632-41.
8. Traeger AC, Henschke N, Hübscher M, Williams CM, Kamper SJ, Maher CG, et al. Estimating the Risk of Chronic Pain: Development and Validation of a Prognostic Model (PICKUP) for Patients with Acute Low Back Pain. PLOS Medicine. 2016;13(5):e1002019.
9. Da Silva T, Macaskill P, Mills K, Maher C, Williams C, Lin C, et al. Predicting recovery in patients with acute low back pain: A Clinical Prediction Model. European Journal of Pain. 2017;21(4):716-26.
10. Von Korff M, Miglioretti DL. A prognostic approach to defining chronic pain. Pain. 2005;117(3):304-13.
11. Wippert P-M, Puschmann A-K, Drießlein D, Arampatzis A, Banzer W, Beck H, et al. Development of a risk stratification and prevention index for stratified care in chronic low back pain. Focus: yellow flags (MiSpEx network). Pain Rep. 2017;2(6):e623-e.
12. van der Gaag WH, Chiarotto A, Heymans MW, Enthoven WTM, van Rijckevorsel-Scheele J, Bierma-Zeinstra SMA, et al. Developing clinical prediction models for non-recovery in older patients seeking care for back pain: the BACE prospective cohort study. PAIN. 2021;162(6):1632-40.
13. Riley RD, Moons KG, Hayden JA, Sauerbrei W, Altman DG. Prognostic Factor Research. In: Riley RD, van der Windt DA, Croft P, Moons KG, editors. Prognosis Research in Healthcare Concepts, Methods and Impact. Oxford, United Kingdom: Oxford University Press; 2019.
14. Steyerberg EW. Clinical prediction models. A practical approach to development, validation, and updating. Second Edition ed. Cham, Switzerland: Springer; 2019.
15. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. BMJ. 2013;346:e5595.
16. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. Annals of Internal Medicine. 2015;162(1):55.
17. Roland M, Morris R. A Study of the Natural History of Back Pain: Part IDevelopment of a Reliable and Sensitive Measure of Disability in Low-Back Pain. spine. 1983;8(2):141-4.
18. Grotle M, Brox J, Vollestad N. Cross-cultural adaptation of the Norwegian versions of the Roland-Morris Disability Questionnaire and the Oswestry Disability Index. Journal of Rehabilitation Medicine. 2003;35(5):241-7.
19. Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. BMJ. 2020:m441.

20. Rundell SD, Sherman KJ, Heagerty PJ, Mock CN, Dettori NJ, Comstock BA, et al. Predictors of Persistent Disability and Back Pain in Older Adults with a New Episode of Care for Back Pain. Pain Med. 2017;18(6):1049-62.
21. Kent P, Lauridsen HH. Managing Missing Scores on the Roland Morris Disability Questionnaire. 2011;36(22):1878-84.
22. Van Buuren S. Flexible imputation of missing data: CRC press; 2018.
23. Riley RD, Moons KG, Debray TPA, Snell KIE, Steyerberg EW, Altman DG, et al. Prognostic Model Research. In: Riley RD, van der Windt DA, Croft P, Moons KG, editors. Prognosis Research in Healthcare: Concepts, Methods, and Impact. Oxford, United Kingdom: Oxfors University Press; 2019. p. 139-87.
24. Debray TPA, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KGM. A new framework to enhance the interpretation of external validation studies of clinical prediction models. Journal of Clinical Epidemiology. 2015;68(3):279-89.
25. Hosmer Jr DW, Lemeshow S, Sturdivant RX. Applied logistic regression: John Wiley & Sons; 2013.
26. Gulia KK, Kumar VM. Sleep disorders in the elderly: a growing challenge. Psychogeriatrics. 2018;18(3):155-65.
27. Jesus-Moraleida FR, Ferreira PH, Ferreira ML, Silva JP, Maher CG, Enthoven WT, et al. Back Complaints in the Elders in Brazil and the Netherlands: a cross-sectional comparison. Age and ageing. 2017;46(3):476-81.
28. Alsaadi SM, McAuley JH, Hush JM, Lo S, Lin CW, Williams CM, et al. Poor sleep quality is strongly associated with subsequent pain intensity in patients with acute low back pain. Arthritis Rheumatol. 2014;66(5):1388-94.
29. Chen Q, Hayman LL, Shmerling RH, Bean JF, Leveille SG. Characteristics of chronic pain associated with sleep difficulty in older adults: the Maintenance of Balance, Independent Living, Intellect, and Zest in the Elderly (MOBILIZE) Boston study. J Am Geriatr Soc. 2011;59(8):1385-92.
30. Blågestad T, Pallesen S, Lunde L-H, Sivertsen B, Nordhus I-H, Grønli J. Sleep in Older Chronic Pain Patients: A Comparative Polysomnographic Study. The Clinical Journal of Pain. 2012;28(4):277-83.

Statistical Analysis Plan (SAP) for:

2-year clinical course of pain intensity and symptom satisfaction for latent classes in older adults with back pain

SAP authors:

Ann-Christin Sannes, Lise Kretz, Are Hugo Pripp, Ørjan Nesse Vigedal, Rikke Munk Killingmo, Silje Stensrud, Kjersti Storheim, Margreth Grotle, Iben Axén.

Project:

BACk pain in Elders in Norway (BACE-N): a prospective cohort study of older people visiting primary care with a new episode of back pain

NCT identifier: NCT04261309

Document date: 21.02.2024

Contents

1. Administrative Information

Introduction to SAP

This document is a supplement to the BACE-N protocol (ClinicalTrials.gov identifier: NCT04261309) and comprises a statistical analysis plan (SAP) for the article with working title "2-year clinical course of pain intensity and symptom satisfaction for latent classes in older adults with back pain".

Study

BACk pain in Elders in Norway (BACE-N): a prospective cohort study of older people in primary care with a new episode of back pain.

Sponsors

Oslo Metropolitan University, the Norwegian Fund for Post-Graduate Training in Physiotherapy, Norway and Et Liv I Bevegelse (A life in movement – Norwegian Chiropractors' Research Foundation).

Approvals

The study was classified as a quality assessment study by the Norwegian Regional Committee for Medical Research Ethics (reference no. 2014/1634/REK vest) and was approved by the Norwegian Social Science Data Service (reference no. 42149) in 2015.

Data storage

Signed informed consent forms are stored in a locked storage. The personal data registry is stored in a secure TSD server, accessible only with a project-specific username, user-specific password and one-time code. Collected, unidentified data is stored in the TSD secure database. Data received through regular mail were manually entered directly into the same database. The original paper questionnaires are stored in a locked room – physically removed from the informed consent forms.

Version of SAP

Version 2.0, 21.02.2024.

Names, affiliations, and roles of SAP contributors

Scientific board:

Chris Maher, Bart Koes, Manuela Ferreira, Kjersti Storheim, Liv Inger Strand, Silje Stensrud, Iben Axén, Thor Einar Holmgaard, Margreth Grotle.

Name	Title	Affiliation	Role
Ann-Christin Sannes	DC, PhD-student	Oslo Metropolitan University	Main author of SAP
Lise Kretz	DC, PhD	Oslo Metropolitan University	Contributor to SAP
Are Hugo Pripp	Professor, PhD	Oslo Metropolitan University	Statistical advisor
Ørjan Nesse Vigdal	PT, PhD-student	Oslo Metropolitan University	Contributor to SAP
Rikke Munk Killingmo	PT, PhD	Oslo Metropolitan University	Contributor to SAP
Silje Stensrud	PT, associate professor, PhD	Oslo Metropolitan University	Contributor to SAP
Kjersti Storheim	PT, professor, PhD	Oslo Metropolitan University, Oslo University Hospital	Contributor to SAP
Margreth Grotle	PT, professor, PhD	Oslo Metropolitan University P44, 0167 Oslo Phone: 90111172 Email: mgrotle@oslomet.no	Principal Investigator
Iben Axén	DC, associate professor, PhD	Karolinska Institutet	Contributor to SAP

Authors

Ann-Christin Sannes, Lise Kretz, Are Hugo Pripp, Ørjan Nesse Vigedal, Rikke Munk Killingmo, Silje Stensrud, Kjersti Storheim, Margreth Grotle, Iben Axén

Abbreviations

BACE	BACk Complaints in the Elderly
BBQ	Back Beliefs Questionnaire
CES-D	Center for Epidemiological Studies Depression questionnaire
FABQA-PA	FearAvoidance Belief Questionnaire – Physical Activity subscale
GLMM	General linear mixed model
LBP	Low Back Pain
LCA	Latent Class Analysis
NRS	Numeric Rating Scale
(PASS)	Patient Acceptable Symptom State
OR	Odds ratio
PCS	Pain Catastrophizing Scale
RMDQ	Roland-Morris Disability Questionnaire
SAP	Statistical Analysis Plan
SCQ	Self-Administered Comorbidity Questionnaire
TSD	Services for Sensitive Data (Tjeneste for Sensitive Data)

Signatures

_____	_____	_____
Ann-Christin Sannes, main author of the SAP	Are Hugo Pripp, senior statisticial advisor	Margreth Grotle principal investigator

2. Background and objectives

Background

Back pain represents a considerable burden on the society as well as the individual (1–6) and most people will experience back pain at some point during their life (7). Among older adults, low back pain (LBP) is the most common health problem that results in considerable disability (8,9). Many people seek primary health care for their back pain, and those aged 65 years or older is the second most common age group to visit physicians for LBP (10).

Back pain is recognized to be an episodic condition, with periods of flare-ups and (partly) recovery (11). The condition is highly heterogenic in its presentation and course, and consequently, challenging for clinicians and patients to predict outcome and to decide on the best available treatment. Further knowledge on the clinical course of back pain in the older patients may provide a better understanding of the variability of recovery. Repeated measurements in cohort studies across different settings and countries have identified common back pain trajectory patterns, reflecting the heterogeneity of back pain course across individuals of different age groups (11–14). Previous evidence shows that for back pain trajectory studies, most studies have identified between 4 and 5 trajectory patterns for back pain: some differing mainly in terms of severity and less in the course patterns, and some differing in the course patterns (i.e., non-parallel and crossing trajectory lines) (for review see (11)). Further, the review showed that all studies identified a pattern of recovery (except one study that included only LBP patients with at least three months duration) and a pattern of persistent severe LBP (except one study excluding patients with unchanging pain due to their analytic approach). Moreover, patterns of more fluctuating pain, i.e., alternating LBP intensity and/or by LBP episodes with periods of no pain were observed (11). However, all included studies in the review excluded back pain patients >65 years of age.

Recently, some information in older adults have presented, with two large prospective cohort studies investigating the course of back pain in older adults. One found that half of the patients reported clinically important improvement in back pain intensity, but only 43% showed clinically important improvement in back-related disability after 2 years (15). 17% reported no back pain and no back pain-related disability at 2 years. The other found that 43% of older adults presenting to primary care with back pain reported themselves as (almost) recovered after 5-years (16). Moreover, what is considered as clinical improvement may not be as important for the patients. Perhaps adding an element of symptom satisfaction ought to be investigated to assess the patient's perspective on their own condition. One such was to investigate this is by using the

Patient Acceptable Symptom State (PASS) (17). Hence, this outcome will also be important to explore further and will be included in the present paper.

This study will build on a latent class analysis (LCA) performed on baseline data (cross-sectional) from this study population (ref/not published yet). The LCA found four classes based on Numeric Rating Scale (NRS), Self-Administered Comorbidity Questionnaire (SCQ), Roland-Morris Disability Questionnaire (RMDQ), Fear-Avoidance Belief Questionnaire (FABQA), Back Beliefs Questionnaire (BBQ), Center for Epidemiological Studies Depression (CES-D), and Pain Catastrophizing Scale (PCS). These four classes were named “The Positive”, “The Fearful”, “The Negative”, and “The Hopeful”. The first class (“The Positive”) had the overall lowest score for reported pain intensity, disability, comorbidities, symptoms of depression and pain catastrophizing. The second class (“The Fearful”) had the highest score on fear avoidance behaviour and comorbidity, with the second highest score on reported pain intensity, disability, catastrophising, and symptoms of depression. The third class (“The Negative”) had the highest score on most of the measurements. Of all the classes they showed the highest pain intensity, disability, symptoms of depression, and pain catastrophising. Moreover, class 3 had the second highest score on both reported comorbidities and fear avoidance behaviour. The fourth and final class (“The Hopeful”) had the second lowest overall score on all measurements. Unlike class one they had somewhat higher scores on pain intensity, disability, fear avoidance behaviour, symptoms of depression, and pain catastrophising.

Exploring the clinical course of back pain intensity and patient acceptable symptom state based on uncovered latent classes may add prognostic value and eventually contribute to better targeting the treatment of these patients.

Objectives

The primary aim is to explore the 2-year clinical course of back pain and symptom satisfaction among older adults seeking primary health care based on classes identified by latent class analysis.

3. Study population, design, and method

Study population and recruitment

Eligible patients for the BACE-N study were all people aged 55 years or older seeking primary care (general practitioner, physiotherapist, or chiropractor) with a new episode of back pain. A new episode was defined as being preceded by 6 months without visiting a primary care provider for similar complaints. Patients were excluded from the study if they had a cognitive impairment

which precluded them from completing the study questionnaires or if they had difficulties speaking and writing Norwegian. Patients who had severe mobility impairments (e.g., were wheelchair bound) were excluded if they could not complete the physical examination. While the study was ongoing, patients received care as usual.

Between April 2015 and February 2020 patients were recruited from general practitioners, physiotherapists, and chiropractors working in Norwegian primary care. Several of these practices were multidisciplinary clinics.

After completion of the initial clinical examination and questionnaire, patients were given follow-up questionnaires through email or postal mail at 3 months, 6 months, 12 months, and 24 months. The participants were informed that they may withdraw from the study at any time without any explanations or consequences.

Study design and setting

This study is using longitudinal data from an observational cohort study with a 2-year follow-up in a Norwegian primary care setting.

BACE-N is part of an international consortium. Details of the BACE-N are provided in the BACE-N protocol (ClinicalTrials.gov identifier: NCT04261309).

Variables

Outcome measure

The first outcome is average pain intensity during the last week (18,19) measured by the Numeric Rating Scale (NRS, 0 = no pain - 10 = worst pain imaginable). NRS is a subjective measure for pain (20), which is found to be preferable for low back pain cases (21).

The second outcome is symptom satisfaction (17). PASS consists of a single item: “How satisfied would you be if your current symptoms were to persist the rest of your life?”, measured using a 5-point likert scale at 3 months, 6 months, and 1- and 2-year follow-up. The categories are “very satisfied”, “somewhat satisfied”, “neither satisfied nor dissatisfied”, “somewhat dissatisfied” and “very dissatisfied”.

Prognostic factor - Latent classes

A latent class analysis, currently in progress, will be used to identify the number of classes intended as prognostic factors in this analysis. The hypothesis is that there may be a difference in the clinical

course of the chosen outcomes depending on the probability of belonging to the different latent classes mentioned above.

Potential confounding factors

In line with the PROGRESS framework and recommendations for type 2 studies (22), we will adjust for confounding factors. Confounding factors included in the analysis, based on previous literature, related to pain intensity and symptom satisfaction are age (12,15,23), gender (12,24), educational level (25,26), employment status (still in paid work), and visited primary health care practitioner (27) (see table 1).

4. Statistical principles

Statistical analysis

All analyses described in this plan are considered *a priori* in that they have been defined in the protocol and/or in this SAP. In case they will be conducted, all post *hoc* analyses will be identified as such in the article. Further, all analyses will be carried out by the main author of this SAP using the STATA (StataCorp, College Station, TX, USA) or R.

To assess for normality of the variables of interest Q-Q plots and histograms will be used. Normally distributed data will be described using mean and standard deviation for continuous variables and counts and percentage for categorical data. Skewed data will be presented with medians and interquartile ranges.

Two separate analyses will be conducted for the two chosen outcomes. To investigate the association a general linear mixed model (GLMM) will be conducted to investigate the clinical course of both outcomes within the latent classes for all follow-up time points. GLMM is a regression-based technique which takes into account the dependency of the repeated observations of each individual, by adjusting for the correlations in estimating a variance for the intercept, a variance for the slope(s) or through different residual matrices (28). Both the unadjusted and the adjusted model will include the main effect of the classes and the follow-up time points, in addition to the interaction between the two (classes x follow-up). The adjusted model will include the selected covariates mentioned in section 3 (other variables). We will use a subject-specific random intercept.

Description of study flow

The flow of participants throughout all follow-up points will be visualized using a flow chart (see figure 1).

Participant characteristics

Baseline characteristics of eligible patients will be presented as shown in table 1.

Handling of missing data

Dataset including missing values will be used for the GLMM analyses, with no imputations for missing data (29).

Model estimation and selection

Model fit of GLMM will be assessed by comparing the log likelihood values and/or model fit statistics as Akaike Information Criterion (AIC) of neighboring [nested] models to determine whether the inclusion of a random slope is necessary and improves the model fit.

The results will be presented as crude and adjusted beta coefficients (B) with 95% confidence intervals (CI) as shown in table 2 and as margin plots with pairwise comparisons.

Sample size

This study contains secondary analyses embedded in the BACE-N study. The BACE-N study included a total of 452 participants which is deemed as an appropriate number of people for such a study (30). The number of subjects included for each time point (baseline, 3-, 6-, 12-, and 24 months) that answered all questions of interest in relation to the pain intensity analysis was 358, 306, 293, 281, and 248, respectively. In regard to the PASS investigation, these numbers for the same timepoints were 365, 295, 279, 272, and 230, respectively.

5. Presentation of the results

Table 1. Patient Characteristics at baseline based on classes identified by LCA (n =)

	Class 1	Class 2	Class 3	Class 4
Age (mean (SD))				
Gender (n (female))				
Education (n (%))				
Less than High School				
High School				

Higher education <4 years	
Higher education > 4 years	
Paid work (n (%))	
Yes	
No	
Primary health care practitioner (n (%))	
General practitioner	
Physiotherapist	
Chiropractor	
Pain intensity (NRS) (mean (SD))	
Symptom satisfaction (PASS) (mean (SD))	

SD – standard deviation, **n** – number, **NRS** – numeric rating scale, **PASS** – patient acceptable symptom state

Table 2. Estimates of fixed effects parameters from linear mixed models. Showing the association between latent classes, pain intensity (NRS) and symptom satisfaction (PASS) over 24 months.

	NRS score			PASS score		
	B	**95% CI**	***P***	**B**	**95% CI**	***P***
Unadjusted model						
Follow-up time points						
Baseline (ref)						
3 months						
6 months						
12 months						
24 months						
LCA class						
The Positive (ref)						
The Fearful						
The Negative						
The Hopeful						
Interaction Time x LCA						
Baseline x The Positive (ref)						
3 months x The Fearful						

3 months x The Negative

3 months x The Hopeful

6 months x The Fearful

6 months x The Negative

6 months x The Hopeful

12 months x The Fearful

12 months x The Negative

12 months x The Hopeful

24 months x The Fearful

24 months x The Negative

24 months x The Hopeful

Full model

Follow-up time points
Baseline (ref)
3 months
6 months
12 months
24 months
LCA class
The Positive (ref)
The Fearful
The Negative
The Hopeful
Interaction Time x LCA
Baseline x The Positive (ref)
3 months x The Fearful
3 months x The Negative
3 months x The Hopeful
6 months x The Fearful
6 months x The Negative
6 months x The Hopeful
12 months x The Fearful
12 months x The Negative
12 months x The Hopeful
24 months x The Fearful
24 months x The Negative
24 months x The Hopeful

Age	
Gender	
Male (ref)	
Female	
Education	
< High school (ref)	
High school	
< 4 years higher education	
> 4 years higher education	
Payed work	
No (ref)	
Yes	
Primary health care practitioner	
General practitioner (ref)	
Physiotherapist	
Chiropractor	

B – regression coefficient, **P** – p-value, **CI** – confidence interval, **NRS** – Numeric Rating Scale, **PASS** – Patient Acceptable Symptom State, bold numbers indicate statistical significance

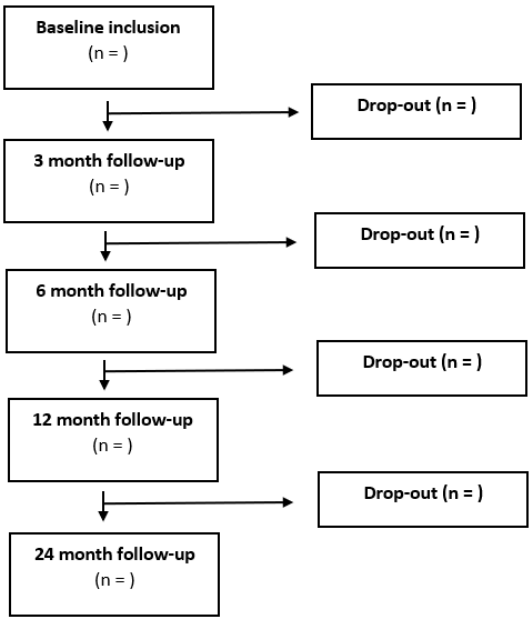


Figure 1a. Flow chart of the study NRS

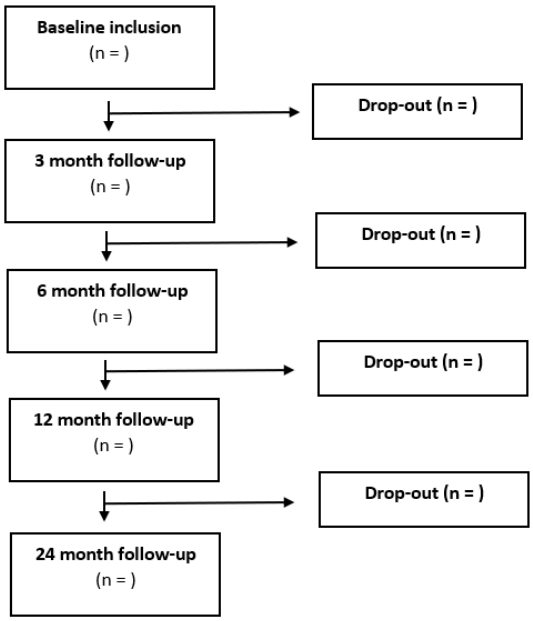


Figure 1b. Flow chart of the study PASS

Figures will be presented to visualize the pain trajectory patterns for the different classes.

6. References

1. Brattberg G, Parker MG, Thorslund M. The prevalence of pain among the oldest old in Sweden. Pain. 1996 Sep;67(1):29–34.

2. Blyth FM, Cumming R, Mitchell P, Wang JJ. Pain and falls in older people. Eur J Pain. 2007 Jul;11(5):564–71.

3. 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. Lancet. 2018 Jun 9;391(10137):2356–67.

4. Ferreira ML, de Luca K. Spinal pain and its impact on older people. Best Pract Res Clin Rheumatol. 2017 Apr;31(2):192–202.

5. Hagen EM, Svensen E, Eriksen HR, Ihlebaek CM, Ursin H. Comorbid subjective health complaints in low back pain. Spine (Phila Pa 1976). 2006 Jun 1;31(13):1491–5.

6. Simon CB, Hicks GE. Paradigm Shift in Geriatric Low Back Pain Management: Integrating Influences, Experiences, and Consequences. Phys Ther. 2018 May 1;98(5):434–46.

7. de Souza IMB, Sakaguchi TF, Yuan SLK, Matsutani LA, do Espírito-Santo A de S, Pereira CA de B, et al. Prevalence of low back pain in the elderly population: a systematic review. Clinics (Sao Paulo). 2019;74:e789.

8. Prince MJ, Wu F, Guo Y, Gutierrez Robledo LM, O'Donnell M, Sullivan R, et al. The burden of disease in older people and implications for health policy and practice. Lancet. 2015 Feb 7;385(9967):549–62.

9. 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 Rheum. 2012 Jun;64(6):2028–37.

10. Wong AY, Karppinen J, Samartzis D. Low back pain in older adults: risk factors, management options and future directions. Scoliosis and Spinal Disorders. 2017 Apr 18;12(1):14.

11. Kongsted A, Kent P, Axen I, Downie AS, Dunn KM. What have we learned from ten years of trajectory research in low back pain? BMC Musculoskelet Disord. 2016 May 21;17:220.

12. Deyo RA, Bryan M, Comstock BA, Turner JA, Heagerty P, Friedly J, et al. Trajectories of symptoms and function in older adults with low back disorders. Spine (Phila Pa 1976). 2015 Sep 1;40(17):1352–62.

13. Enthoven WTM, Koes BW, Bierma-Zeinstra SMA, Bueving HJ, Bohnen AM, Peul WC, et al. Defining trajectories in older adults with back pain presenting in general practice. Age Ageing. 2016 Nov;45(6):878–83.

14. Axén I, Leboeuf-Yde C. Trajectories of low back pain. Best Pract Res Clin Rheumatol. 2013 Oct;27(5):601–12.

15. Jarvik JG, Gold LS, Tan K, Friedly JL, Nedeljkovic SS, Comstock BA, et al. Long-term outcomes of a large, prospective observational cohort of older adults with back pain. Spine J. 2018 Sep;18(9):1540–51.

16. Gaag WH van der, Enthoven WTM, Luijsterburg PAJ, Rijckevorsel-Scheele J van, Bierma-Zeinstra SMA, Bohnen AM, et al. Natural History of Back Pain in Older Adults over Five Years. J Am Board Fam Med. 2019 Nov 1;32(6):781–9.

17. Tubach F, Wells GA, Ravaud P, Dougados M. Minimal clinically important difference, low disease activity state, and patient acceptable symptom state: methodological issues. J Rheumatol. 2005 Oct;32(10):2025–9.

18. Artus M, Campbell P, Mallen CD, Dunn KM, van der Windt DAW. Generic prognostic factors for musculoskeletal pain in primary care: a systematic review. BMJ Open. 2017 Jan 17;7(1):e012901.

19. Hestbaek L, Munck A, Hartvigsen L, Jarbøl DE, Søndergaard J, Kongsted A. Low Back Pain in Primary Care: A Description of 1250 Patients with Low Back Pain in Danish General and Chiropractic Practice. Int J Family Med. 2014;2014:106102.

20. Alghadir AH, Anwer S, Iqbal A, Iqbal ZA. Test–retest reliability, validity, and minimum detectable change of visual analog, numerical rating, and verbal rating scales for measurement of osteoarthritic knee pain. J Pain Res. 2018 Apr 26;11:851–6.

21. Strong J, Ashton R, Chant D. Pain intensity measurement in chronic low back pain. Clin J Pain. 1991 Sep;7(3):209–18.

22. Riley RD, Hayden JA, Steyerberg EW, Moons KGM, Abrams K, Kyzas PA, et al. Prognosis Research Strategy (PROGRESS) 2: Prognostic Factor Research. PLoS Med. 2013 Feb 5;10(2):e1001380.

23. Leopoldino AAO, Megale RZ, Diz JBM, Moreira B de S, Lustosa LP, Pereira LSM, et al. Influence of the number and severity of comorbidities in the course of acute non-specific low back pain in older adults: longitudinal results from the Back Complaints in the Elders (BACE-Brazil). Age Ageing. 2019 Dec 1;49(1):96–101.

24. Jarvik JG, Comstock BA, Heagerty PJ, Turner JA, Sullivan SD, Shi X, et al. Back pain in seniors: the Back pain Outcomes using Longitudinal Data (BOLD) cohort baseline data. BMC Musculoskelet Disord. 2014 Apr 23;15:134.

25. Lim KL, Jacobs P, Klarenbach S. A Population-Based Analysis of Healthcare Utilization of Persons With Back Disorders: Results From the Canadian Community Health Survey 2000–2001. Spine. 2006 Jan 15;31(2):212.

26. Hoebel J, Rattay P, Prütz F, Rommel A, Lampert T. Socioeconomic Status and Use of Outpatient Medical Care: The Case of Germany. PLOS ONE. 2016 May 27;11(5):e0155982.

27. Fritz JM, Kim J, Dorius J. Importance of the type of provider seen to begin health care for a new episode low back pain: associations with future utilization and costs. J Eval Clin Pract. 2016 Apr;22(2):247–52.

28. Hoekstra T, Twisk JWR. The Analysis of Individual Health Trajectories Across the Life Course: Latent Class Growth Models Versus Mixed Models. In: Burton-Jeangros C, Cullati S, Sacker A, Blane D, editors. A Life Course Perspective on Health Trajectories and Transitions [Internet]. Cham (CH): Springer; 2015 [cited 2023 Apr 21]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK385363/

29. Twisk J. Applied Mixed Model Analysis: A Practical Guide. 2019.

30. Chi YY, Glueck DH, Muller KE. Power and Sample Size for Fixed-Effects Inference in Reversible Linear Mixed Models. Am Stat. 2019;73(4):350–9.

Statistical Analysis Plan (SAP) for:

Development of a clinical prediction model for unfavourable patient acceptable symptom state (PASS) in older patients seeking care in Norwegian primary health care services for back pain

SAP authors:

Ann-Christin Sannes, Ørjan Nesse Vigdal, Are Hugo Pripp, Lise Kretz, Rikke Munk Killingmo, Silje Stensrud, Iben Axén, Kjersti Storheim, Margreth Grotle

Project:

BACk pain in Elders in Norway (BACE-N): a prospective cohort study of older people visiting primary care with a new episode of back pain

NCT identifier: NCT04261309

Document date: 15.09.2023

Contents

1. Administrative Information

Introduction to SAP

This document is a supplement to the BACE-N protocol (ClinicalTrials.gov identifier: NCT04261309) and comprises a statistical analysis plan (SAP) for the article with working title "Development of a clinical prediction model for unfavourable patient acceptable symptom state (PASS) in older patients seeking care in Norwegian primary health care services for back pain".

Study

BACk pain in Elders in Norway (BACE-N): a prospective cohort study of older people in primary care with a new episode of back pain.

Sponsors

Oslo Metropolitan University, the Norwegian Fund for Post-Graduate Training in Physiotherapy, Norway, and Et Liv I Bevegelse (A life in movement – Norwegian Chiropractors' Research Foundation).

Approvals

The study was classified as a quality assessment study by the Norwegian Regional Committee for Medical Research Ethics (reference no. 2014/1634/REK vest) and was approved by the Norwegian Social Science Data Service (reference no. 42149) in 2015.

Version of SAP

Version 1.0, 08.07.2023.

Names, affiliations, and roles of SAP contributors

Scientific board:

Chris Maher, Bart Koes, Manuela Ferreira, Kjersti Storheim, Liv Inger Strand, Silje Stensrud, Iben Axén, Thor Einar Holmgaard, Margreth Grotle.

Name	Title	Affiliation	Role
Ann-Christin Sannes	DC, PhD	Oslo Metropolitan University	Main author of SAP
Ørjan Nesse Vigdal	PT, PhD-student	Oslo Metropolitan University	Contributor to SAP

Are Hugo Pripp	Professor, PhD	Oslo Metropolitan University	Statistical advisor
Lise Kretz	DC, PhD-student	Oslo Metropolitan University	Contributor to SAP
Rikke Munk Killingmo	PT, PhD	Oslo Metropolitan University	Contributor to SAP
Silje Stensrud	PT, associate professor, PhD	Oslo Metropolitan University	Contributor to SAP
Iben Axén	DC, associate professor, PhD	Karolinska Institutet	Contributor to SAP
Kjersti Storheim	PT, professor, PhD	Oslo Metropolitan University, Oslo University Hospital	Contributor to SAP
Margreth Grotle	PT, professor, PhD	Oslo Metropolitan University P44, 0167 Oslo Phone: 90111172 Email: mgrotle@oslomet.no	Principal Investigator

Authors

Ann-Christin Sannes, Ørjan Nesse Vigdal, Are Hugo Pripp, Lise Kretz, Rikke Munk Killingmo, Silje Stensrud, Iben Axén, Kjersti Storheim, Margreth Grotle

Abbreviations

AIC	Akaike information criteria
AUC	Area under Receiver Operator Characteristics Curve
BACE	BACk Complaints in the Elderly
CES-D	Center for Epidemiological Studies Depression questionnaire
CITL	Calibration-in-the-large
FABQ-PA	Fear-Avoidance Belief Questionnaire – Physical Activity subscale
KOOS	Knee injury and Osteoarthritis Outcome Score
BP	Back Pain
NRS	Numeric Rating Scale
PASS	Patient Acceptable Symptom State
PCS	Pain Catastrophizing Scale
RMDQ	Roland-Morris Disability Questionnaire
SAP	Statistical Analysis Plan
SCQ	Self-Administered Comorbidity Questionnaire
TRIPOD	Transparent Reporting of a multivariable prediction model for the Individual Prognosis Or Diagnosis
TSD	Services for Sensitive Data (Tjeneste for Sensitive Data)

Signatures

Ann-Christin Sannes, main author of the SAP	Are Hugo Pripp, senior statistician responsible	Margreth Grotle principal investigator

2. Background and objectives

Background

Back pain represents a considerable burden on the society as well as the individual (1–6) which effects most people at some point (7). However, back pain has an increased prevalence and unfavorable consequences among older adults (8,9). Since persistent, or chronic, pain is difficult to successfully treat, including other aspects, such as patient-based evaluation of their own symptoms, ought to be included in pain research. One measurement that has gained some traction in the recent years is the patient acceptable symptom state (PASS) item (10,11). This may be an important tool to monitor the impact symptoms may have on a person's daily life which could be helpful in a treatment management setting.

Another important part of patient management is the use of prognostic models to present a possible individual prognosis for the patients. This is done by considering several prognostic factors that may be important for accurate prediction (12). The use of prognostic models has yet to be extensively investigated in the older population suffering from back pain (13), and has, to the best of our knowledge, not been assessed in this patient group in relation to PASS until now.

Study aim

The primary aim of this study is to create a prediction model for patient acceptable symptom state (PASS) among older people seeking primary care for a new episode of back pain over a 24-month period. The secondary aim, if deemed appropriate, is to create a clinically useful risk score scale based on the most important risk factors from the prediction model.

Here we present the Statistical Analysis Plan (SAP) for the intended study, with the aim of reducing the risk of data selection bias and data-driven interpretation of results.

3. Study population, design, and method

Study population and recruitment

Eligible patients for the BACE-N study were people aged 55 years or older seeking primary care (general practitioner, physiotherapist, or chiropractor) with a new episode of back pain. A new episode was defined as being preceded by 6 months without visiting a primary care provider for similar complaints. Patients were excluded from the study if they had a cognitive impairment which precluded them from completing the study questionnaires or if they had difficulties speaking and writing Norwegian. Patients who had severe mobility impairments (e.g., were wheelchair bound) were excluded if they could not complete the physical examination. While the study was ongoing, patients received care as usual.

Between April 2015 and February 2020 patients were recruited from general practitioners, physiotherapists, and chiropractors working in Norwegian primary care. Several of these practices were multidisciplinary clinics.

After completion an initial clinical examination and questionnaire, patients were given follow-up questionnaires through email or postal mail at 3 months, 6 months, 12 months, and 24 months. The participants were informed that they may withdraw from the study at any time without any explanations or consequences.

Study design

BACE-N is part of an international consortium, the Back Complaints in the Elders (BACE) (14). Details of the BACE-N are provided in the BACE-N protocol (ClinicalTrials.gov identifier: NCT04261309).

The BACE-N study is a prospective observational cohort study with a 2-year follow-up in a Norwegian primary care setting. This article will use data from baseline, 12 and 24-months. The design of the current study is informed by the framework PROGnosis RESearch Strategy (PROGRESS) type 3 (12). Reporting will be according to the Transparent Reporting of a multivariable prediction model for the Individual Prognosis Or Diagnosis (TRIPOD) statement prediction study (15).

Variables

Outcome measure

The outcome is the patient acceptable symptom state (PASS). PASS consists of a single item: "How satisfied would you be if your current symptoms were to persist the rest of your life?", measured using a 5-point likert scale at 1- and 2-year follow-up. The PASS scores is 1) very satisfied, 2) somewhat satisfied, 3) neither satisfied nor dissatisfied, 4) somewhat dissatisfied, and 5) very dissatisfied. A dichotomization will be done between 2) and 3) where those scoring ≥3 at both 12 and 24 months, will be considered unfavorable.

Prognostic factor

A prediction model created by the Dutch arm of the BACE consortium (13), which was externally validated by our project group (16) investigating pain and disability assessed 14 factors as potential prognostic factors. The same factors, in addition to PASS, will be utilized in this work. This is due to the overlap of some of the same prognostic factors such as pain (13,17,18) and PASS (19), and seems clinically reasonable to investigate.

Table 1. Potential prognostic factors.

Prognostic factor	**Measurement level**	**Measured with**	**Parameters in model**
Age	Continuous		1
Sex	Binary		1
Chronic duration of current episode of back pain	Binary	Duration >3 months at baseline	1
Back pain intensity past week	Continuous, 0-10	NRS	1
Back-related disability	Continuous, 0-24	RMDQ	1
Recent episode of back pain	Binary	An episode of back pain within the past 6 months, for which they did not seek care	1
Comorbidity	Binary	SCQ	1
Radiating pain in the leg	Binary	Presence of radiating leg pain in either or both leg(s)	1
Spinal morning stiffness > 60 minutes	Binary	Modified KOOS-question 6a	1
Pain during spinal rotation	Binary	Pain with rotation of the upper body during clinical examination	1

Expectations of recovery	Binary	Likert scale ranging from 1, "fully recovered" to 5, "worse than ever". Dichotomised into "expecting improvement" and "not expecting improvement".	1
Depressive symptomology	Continuous, 0-60	CES-D	1
Kinesiophobia	Continuous, 0-28	FABQ-PA	1
Pain catastrophizing	Continuous, 0-52	PCS	1
Patient acceptable symptom state	Ordinal, 1-5	PASS	1
NRS; Numeric Rating Scale, RMDQ; Roland-Morris Disability Questionnaire, SCQ; Self-administered Comorbidity Questionnaire, KOOS; Knee injury and Osteoarthritis Outcome Score, CES-D; Center for Epidemiological Studies – Depression, FABQ-PA; Fear-Avoidance Beliefs Questionnaire – Physical Activity subscale, PCS; Pain catastrophizing Scale, PASS; Patient Acceptable Symptom State.			

4. Statistical analyses

General statistical considerations

All analyses described in this plan are considered *a priori* in that they have been defined in the protocol and/or in this SAP. In case they will be conducted, all post *hoc* analyses will be identified as such in the article. Further, all analyses will be carried out by the main author of this SAP using the software STATA (StataCorp, College Station, TX, USA), under supervision from the principal investigator (MG) and the advisory statistician (AHP). The statistical tests will be 2-sided, and p-value less than 0.05 will be considered statistically significant. 95% confidence intervals will be reported on effect sizes.

To assess for normality of the variables of interest Q-Q plots and histograms will be used.

Sample size

The BACE-N cohort consists of a total of 452 participants. To assess appropriate sample size an online sample size calculator available at https://riskcalc.org/pmsamplesize/ was used, as recommended by Riley et al (21). Further, previous evidence has proposed that 100-200 events will be sufficient (22) for the intended analyses.

Handling of missing data

Missing data for prognostic factors at baseline and PASS at item level at 12-, and 24-months will be handled through multiple imputation if necessary (>5% missing), as this may have an impact on the final model. The imputation will then be based on the variable's original scale before dichotomization and combining the time points.

Descriptive statistics

Descriptive data will be reported using mean and standard deviation for normally distributed continuous variables, median and interquartile range for continuous variables with skewed distribution, and with frequency and proportions for categorical variables (Table 2).

Backward selection stepwise regression analyses

To build the most optimal regression model a stepwise selection approach will be used based on the previously validated model. This will be done in three steps. Step 1: fit a regression model using all variables considered prognostic factors and calculate the Akaike information criterion (AIC). Step 2: remove the factor that has the least significant p-value or causes the lowest drop in R^2. Step 2 will be repeated until removing any prognostic factors no longer affects or increases the AIC. The stopping rule chosen for the purpose of the present paper will be p-value <0.154. A graphical comparison will be conducted to assess for linearity for continuous variables. To assess the effect of possible non-linearity of continuous variables, a fractional polynomial regression will be conducted for comparison.

Further, a sensitivity analysis will be conducted based on other variable selection methods as stepwise regression using other criteria or penalized regression.

Results from backward selection stepwise regression analyses will be presented as shown in table 4 for the number of prognostic factors found to be of importance.

Model performance

Model performance will be measured by calibration-in-the-large (CITL), calibration slope, Nagelkerke`s pseudo-$R^2$, and Area under Receiver Operator Characteristics Curve (AUC), as shown in table 3. CITL and calibration slope will be assessed. CITL value equal to 0 and a calibration slope of 1 indicates perfect calibration (23). Nagelkerke`s pseudo-R^2 measures overall model performance that ranges between 0 (worst possible performance) and 1 (best possible performance). AUC measures discriminative performance. Ranging from 0 to 1, where 1 is perfect discrimination and 0.5 indicates discrimination no better than chance. AUC values between 0.9-1.0 is considered outstanding, 0.8-0.9 is excellent and 0.7-0.8 is considered acceptable (24).

Internal validation will be conducted by bootstrapping samples to correct for overoptimism (20). The estimated slope value will be used as a shrinkage factor that will be multiplied with the pooled coefficients and a new intercept will be determined that aligns with the shrunken coefficients. The adjusted linear predictors will be reported in the penalized models and each adjusted models` performance will be evaluated with an optimism-adjusted R^2 and AUC value.

Risk score scale

A risk score scale is a scale that is intended to indicate the risk of certain outcomes in a clinical setting. The scale is meant to be quick and easy to use for the healthcare practitioner, which can aid in the decision and planning of further treatment.

The possibility of creating a risk score scale from the results of the regression analyses will be considered if deemed clinically appropriate.

5. Presentation of the results

Below are suggested drafts of tables used for presentation of the results.

Table 2. Patient Characteristics at baseline

	All participants (n = 452)	Missing, n (%)
Age, mean (SD)		
Female, n (%)		
One or more comorbidities, n (%)		
Recent episode of back pain, n (%)		
Duration of current episode >3 months, n (%)		
Radiating leg pain, n (%)		
Spinal morning stiffness > 60 minutes[a], n (%)		
Pain with active rotation of the back, n (%)		
Average back pain intensity last week (NRS, 0-10), mean (SD)		
Disability (RMDQ, 0-24), mean (SD)		
Duration of current episode >3 months, n (%)		
Kinesiophobia (FABQ-PA, 0-24), mean (SD)		
Depression (CES-D, 0-60), mean (SD)		
Pain catastrophizing (PCS, 0-52), mean (SD)		
Expectations of being fully recovered or much better from back pain within the next 3 months, n (%)		
Symptom satisfaction (PASS, 1-5), mean (SD)		

SD – standard deviation, **n** – number, **SF-36** – Short Form health survey 36, **NRS** – numeric rating scale, **RMDQ** – Roland-Morris Disability Questionnaire, **FABQ-PA** – Fear-Avoidance Beliefs Questionnaire - Physical Activity, **CES-D** – Center for Epidemiological Studies - Depression, **PCS** – Pain Catastrophising scale, **PASS** – patient acceptable symptom state

Table 3. Performance measure for current model		
Persistent dissatisfaction (≥3/10 PASS at both 12 and 24 months)		
Aspect	*Measure*	*Value*
Overall performance	Nagelkerke pseudo-R^2	
Discrimination	Area under the curve	
Calibration	Calibration in-the-large	
	Calibration slope	

Table 4. The results of stepwise backward regression for dependent variable PASS					
R = , R^2 = , Adjust. R^2 = , F() = p					
	Beta Coef	**Standard error**	**Directional coefficient b**	**Standard error with b**	**p-value**
Intercept					
Prognostic factor 1					
Prognostic factor 2					
Prognostic factor 3					

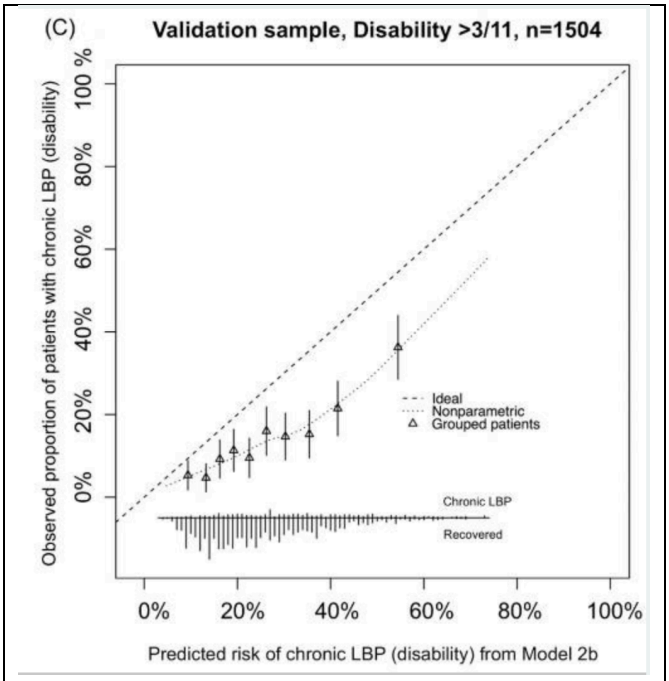


Figure 1. Calibration plot (25).

A calibration plot like figure 1 will be presented to visualize performance measure of current model.

6. References

1. Brattberg G, Parker MG, Thorslund M. The prevalence of pain among the oldest old in Sweden. Pain. 1996 Sep;67(1):29–34.

2. Blyth FM, Cumming R, Mitchell P, Wang JJ. Pain and falls in older people. Eur J Pain. 2007 Jul;11(5):564–71.

3. 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. Lancet. 2018 Jun 9;391(10137):2356–67.

4. Ferreira ML, de Luca K. Spinal pain and its impact on older people. Best Pract Res Clin Rheumatol. 2017 Apr;31(2):192–202.

5. Hagen EM, Svensen E, Eriksen HR, Ihlebaek CM, Ursin H. Comorbid subjective health complaints in low back pain. Spine (Phila Pa 1976). 2006 Jun 1;31(13):1491–5.

6. Simon CB, Hicks GE. Paradigm Shift in Geriatric Low Back Pain Management: Integrating Influences, Experiences, and Consequences. Phys Ther. 2018 May 1;98(5):434–46.

7. de Souza IMB, Sakaguchi TF, Yuan SLK, Matsutani LA, do Espírito-Santo A de S, Pereira CA de B, et al. Prevalence of low back pain in the elderly population: a systematic review. Clinics (Sao Paulo). 2019;74:e789.

8. Prince MJ, Wu F, Guo Y, Gutierrez Robledo LM, O'Donnell M, Sullivan R, et al. The burden of disease in older people and implications for health policy and practice. Lancet. 2015 Feb 7;385(9967):549–62.

9. 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 Rheum. 2012 Jun;64(6):2028–37.

10. Maksymowych WP, Richardson R, Mallon C, van der Heijde D, Boonen A. Evaluation and validation of the patient acceptable symptom state (PASS) in patients with ankylosing spondylitis. Arthritis Care & Research. 2007;57(1):133–9.

11. Salaffi F, Carotti M, Gutierrez M, Di Carlo M, Angelis R. SAT0080 Patient Acceptable Symptom State (PASS) in Self-Report Questionnaires and Composite Clinical Disease Index for Assessing Rheumatoid Arthritis. Identification of Cut-Off Points for Routine Care: Table 1. BioMed Research International. 2015 Jun 18;2015:8.

12. Steyerberg EW, Moons KGM, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10(2):e1001381.

13. van der Gaag WH, Chiarotto A, Heymans MW, Enthoven WTM, van Rijckevorsel-Scheele J, Bierma-Zeinstra SMA, et al. Developing clinical prediction models for nonrecovery in older patients seeking care for back pain: the back complaints in the elders prospective cohort study. Pain. 2021 Jun 1;162(6):1632–40.

14. Scheele J, Luijsterburg PA, Ferreira ML, Maher CG, Pereira L, Peul WC, et al. Back Complaints in the Elders (BACE); design of cohort studies in primary care: an international consortium. BMC Musculoskeletal Disorders. 2011 Aug 19;12(1):193.

15. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. BMC Medicine. 2015 Jan 6;13(1):1.

16. Vigdal ØN, Storheim K, Killingmo RM, Rysstad T, Pripp AH, van der Gaag W, et al. External validation and updating of prognostic prediction models for nonrecovery among older adults seeking primary care for back pain. PAIN. 2022 May 13;10.1097/j.pain.0000000000002974.

17. Rundell SD, Sherman KJ, Heagerty PJ, Mock CN, Dettori NJ, Comstock BA, et al. Predictors of Persistent Disability and Back Pain in Older Adults with a New Episode of Care for Back Pain. Pain Med. 2017 Jun 1;18(6):1049–62.

18. Campbell P, Foster NE, Thomas E, Dunn KM. Prognostic indicators of low back pain in primary care: five-year prospective study. J Pain. 2013 Aug;14(8):873–83.

19. Pårup L, Andersen NB de V, Christiansen DH. Modifiable predictors for acceptable symptom state and impression of change among musculoskeletal disorders in physical therapy patients: a prospective cohort study.

20. Fernandez-Felix BM, García-Esquinas E, Muriel A, Royuela A, Zamora J. Bootstrap internal validation command for predictive logistic regression models. The Stata Journal. 2021 Jun 1;21(2):498–509.

21. Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. BMJ. 2020 Mar 18;368:m441.

22. Riley RD, Moons KG, Hayden JA, Sauerbrei W, Altman DG. Prognostic factor research. In: van der Windt DA, Hemingway H, Croft P, Riley RD, van der Windt D, Croft P, et al., editors. Prognosis Research in Healthcare: Concepts, Methods, and Impact [Internet]. Oxford University Press; 2019 [cited 2023 Jun 22]. p. 0. Available from: https://doi.org/10.1093/med/9780198796619.003.0007

23. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating [Internet]. Cham: Springer International Publishing; 2019 [cited 2023 Jun 23]. (Statistics for Biology and Health). Available from: http://link.springer.com/10.1007/978-3-030-16399-0

24. Application of Logistic Regression with Different Sampling Models. In: Applied Logistic Regression [Internet]. John Wiley & Sons, Ltd; 2013 [cited 2023 Jun 23]. p. 227–42. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118548387.ch6

25. Traeger AC, Henschke N, Hübscher M, Williams CM, Kamper SJ, Maher CG, et al. Estimating the Risk of Chronic Pain: Development and Validation of a Prognostic Model (PICKUP) for Patients with Acute Low Back Pain. PLoS Med. 2016 May 17;13(5):e1002019.