

BMJ Open Independent medical evaluation of general practitioners' follow-up of sick-listed patients: a cross-sectional study in Norway

Irene Øyeflaten ^{1,2}, Silje Maeland,² Inger Haukenes²

To cite: Øyeflaten I, Maeland S, Haukenes I. Independent medical evaluation of general practitioners' follow-up of sick-listed patients: a cross-sectional study in Norway. *BMJ Open* 2020;**10**:e032776. doi:10.1136/bmjopen-2019-032776

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-032776>).

Received 05 July 2019

Revised 13 February 2020

Accepted 26 February 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Norwegian National Advisory Unit on Occupational Rehabilitation, Rauland, Norway

²NORCE Norwegian Research Centre AS, Bergen, Norway

Correspondence to

Dr Irene Øyeflaten;
irene.oyeflaten@arbeidoghelse.no

ABSTRACT

Objectives The study was designed to examine the sufficiency of general practitioners' (GPs) follow-up of patients on sick leave, assessed by independent medical evaluators.

Design Cross-sectional study

Setting Primary health care in the Western part of Norway. The study reuses data from a randomised controlled trial—the Norwegian independent medical evaluation trial (NIME trial).

Participants The intervention group in the NIME trial: Sick-listed workers having undergone an independent medical evaluation by an experienced GP at 6 months of unremitting sick leave (n=937; 57% women). In the current study, the participants were distributed into six exposure groups defined by gender and main sick leave diagnoses (women/musculoskeletal, men/musculoskeletal, women/mental, men/mental, women/all other diagnoses and men/all other diagnoses).

Outcome measure The independent medical evaluators assessment (yes/no) of the sufficiency of the regular GPs follow-up of their sick-listed patients.

Results Estimates from generalised linear models demonstrate a robust association between men with mental sick leave diagnoses and insufficient follow-up by their regular GP first 6 months of sick leave (adjusted relative risk (RR)=1.8, 95% CI=1.15–1.68). Compared with the reference group, women with musculoskeletal sick leave diagnoses, this was the only significant finding. Men with musculoskeletal diagnoses (adjusted RR=1.4, 95% CI=0.92–2.09); men with other diagnoses (adjusted RR=1.0, 95% CI=0.58–1.73); women with mental diagnoses (adjusted RR=1.2, 95% CI=0.75–1.77) and women with other diagnoses (adjusted RR=1.3, 95% CI=0.58–1.73).

Conclusions Assessment by an independent medical evaluator showed that men with mental sick leave diagnoses may be at risk of insufficient follow-up by their GP. Efforts should be made to clarify unmet needs to initiate relevant actions in healthcare and work life. Avoiding marginalisation in work life is of the utmost importance.

Trial registration number NCT02524392; Post-results.

BACKGROUND

General practitioners (GPs) worldwide report conflicting roles in the area of work and health. A well-known conflict within insurance medicine is acting as the patient's

Strengths and limitations of this study

- Reuse of data from the intervention group in a randomised controlled trial, ensured high generalisability of the findings to sick-listed workers in Norway.
- The independent medical evaluators who provided the outcome measure (sufficient follow-up or not), were experienced general practitioners with specialised training in evaluating follow-up of sick-listed workers.
- Data on sick leave diagnoses were provided from the Norwegian national register of social security, securing precise distribution into exposure groups.
- Non-participation among invited may cause a selection bias; however, underestimation of risk is more likely than overestimation.
- One might question the reliability of the sick leave diagnosis since long-term disability is associated with both medical and non-medical factors.

advocate or the society's gatekeeper in sick leave decisions. GPs value higher and feel more competent in treating patients as opposed to guarding the gates to the welfare state's social security schemes.^{1 2} According to the Handbook of Work Disability, healthcare providers are expected to empower their sick-listed patients to take responsibility of their own health in relation to work disability prevention.³ GPs describe challenges in this intersection between work and health, particularly in assessing functional ability, work ability and percentage of work-capacity.^{1 2} Norwegian GPs are important stakeholders in identifying and understanding barriers and facilitators of return to work (RTW) because they manage about 80% of workers on sick leave.⁴ Subsequently, authorities urge GPs to provide sufficient follow-up actions aiming at RTW² and to limit the length of sick leave. Independent medical evaluations (IMEs) are used in different jurisdictions as a second opinion to determine the functional ability of workers who claim the inability to work

due to illness or injury.⁵ The general hypothesis is that an IME eliminate the potential bias of a long-term patient-physician relationship.

European trends demonstrate higher sick leave rates among women compared with men, and this remains largely unexplained in the literature.^{6–9} However, several studies have found gender bias, that is, an unintended and systematic neglect of women during medical examinations, decision-making, treatment and follow-up.^{10–16} Whether a one-off IME consultation means a lower risk of gender bias for sick-listed patients, has not previously been examined.

Patient with musculoskeletal disorders (MSD) and common mental disorders (CMD) account for the majority of sickness absence and disability pension across Europe.¹⁷ The diagnoses are used to operationalise medical reasons for inability to work. However, an overweight of sick leave diagnoses, such as musculoskeletal and common mental, is based on the patients' perception of their health problems with lack of biological correlates. These conditions have low status among professionals compared with more acute and dramatic conditions involving heart, cancer, and internal medicine.¹⁸ Hence, the risk of managing these patients with suboptimal insight and interest is present. Finally, MSD and CMD are more prevalent among women.¹⁹ The gendered distribution combined with the low status of these disorders may influence the follow-up towards women in particular. Thus, intersections of gender and diagnoses have the potential to activate practices and decisions among GPs that cause unwarranted variation in the follow-up of sick-listed patients.

The aim of this study was to use IMEs evaluation to examine the sufficiency of GPs follow-up actions, across intersectional patient groups defined by gender and major sick leave diagnoses, within 6 months of sick leave.

MATERIALS AND METHODS

Setting

This cross-sectional study is based on data from the Norwegian IME trial (NIME trial), a randomised controlled trial conducted in a primary healthcare setting in the Western part of Norway.⁵ The Norwegian Ministry of Labour and Social Affairs initiated and funded the NIME trial. The trial aimed to investigate if new evaluations made by an independent, experienced and specially trained GP (IME physicians) contributed to new perspectives on the actual sick leave compared with evaluations made by the patients regular GP²⁰ and whether sufficient follow-up actions were fulfilled. The training included in-depth knowledge of specific follow-up actions aiming at RTW available in Norway in 2015 and 2016.^{5 21} The rationale for an IME was to eliminate the potential bias of a long-term patient-physician relationship.⁵ The IME physicians were employed by the Norwegian Labour and Welfare Administration (NAV). As employed by NAV, the IME physician

had access to a short summary from the regular GP and all recorded follow-up actions past 6 months.

The IME physicians had 2 hours available per patient, whereof 1 hour was earmarked the patient consultation. The consultation followed a predefined scheme resulting in a report (the IME report) about the patients' background and opinions on different issues concerning their sickness absence, prior actions aiming at RTW and finally recommendations of further sick leave level and evaluation of the sufficiency of the prior actions. The IME physicians' report was sent to the regular GP with peer support suggestions for further follow-up of the sick-listed worker.

Population

The NIME trial included all workers (except those with dementia, cancer, and pregnancy-related diagnoses) with unremitting sick leave (full or partial) the last 6 months in Hordaland county, Norway, between March 2015 and March 2016 (n=5888). A block randomisation (1:1) was performed determined by a computer-generated randomisation list. The protocol stated that workers randomised to an IME consultation (n=2616) should receive a letter from the local NAV office with information and time and place for the consultation. The 2599 participants randomised to treatment as usual did not receive any information and were merely followed-up as usual by their regular GP and NAV. Among those who were randomised to an IME consultation, 918 workers did not receive an invitation due to capacity problems. Among those receiving the invitation, 761 workers cancelled or did not show up for the appointment. The current study included those who were randomised to the IME and exposed to the intervention (n=937). Potential selection bias related to those who did not participate was addressed by comparing relevant characteristics (gender, age and sick leave diagnosis) between the groups.²²

Data sources

Data sources were the IME reports and a national register of sickness benefits from NAV. Information about age, country of birth, gender, marital status, income last year, occupation, sick leave diagnosis and days on sick leave during the last year (before the IME) were register-based (NAV). Information about 'follow-up actions aiming at RTW' came from the IME reports. The data sources were linked through the unique personal ID number given to all Norwegian inhabitants.

Exposure

GP certified sickness absence in Norway requires a diagnosis from The International Classification of Primary Care.²³ To combine gender with major sick leave diagnoses we first established three groups (1) chapter L=MSD, (2) chapter p=mental diagnoses and (3) all other diagnoses. Second, we combined gender with each of these three groups generating six exposure groups with sick leave diagnoses (women/MSD, men/MSD, women/mental, men/mental, women/other and men/other).

Outcomes

Information about whether *sufficient follow-up actions* was carried out or not by the regular GP during the 6 months prior to the IME-consultation, was based on the IME physicians' evaluations and obtained from the IME-report. The IME physicians answered yes or no on the following question: 'Do you think that adequate and appropriate follow-up actions have been taken?'

Covariates

Age was registered as a continuous variable, and *marital status* as (married or not, single or not, divorced or not). *Country of birth* was registered in country codes and recoded into three categories (Norway, Western and Non-Western). *Income last year* was registered in Norwegian kroneNOK and categorised in percentiles (NOK \geq 517644, NOK 422 448 to 517643, NOK 329 402 to 422 447 and NOK <329 401). *Occupation* was registered according to Standard Classification of Occupations (ISCO-codes).²⁴ First, we generated nine main occupational categories by using the first number in the code-structure. Thereafter, we recoded the nine categories into four categories (managers/professionals, associate professionals, routine non-manual and manual).^{25–27} *Sickness absence last 12 months* (pre-baseline) was registered in days.

Statistical analyses

We used Pearson χ^2 and t-test to examine gender differences in the distributions and means across covariates. We estimated univariable associations between covariates and outcomes using generalised linear models (GLM).²⁸ Selection of variables to control for potential confounding in the adjusted GLM was based on recommendations from Talbot and Massambla²⁹ and Hernán *et al.*³⁰ We used a combination of statistical associations from the data and prior knowledge about the associations that links exposure, outcome and potential confounders. The reference group for the exposure variable was women/MSD, and for variables in the adjusted GLM: Norway (country of birth), managers/professionals (occupation). The GLM procedure with binomial distribution and log link, generates estimates presented as relative risks (RRs) with 95% CIs. Missing items on variables were substituted with information from the IME report, whenever possible.

Patient and Public Involvement

Patients and public were not involved in the design or planning of the current study.

RESULTS

A majority of the participants were women (57%). Mean age among men and women were 46 and 47 years, respectively, with corresponding SDs of 11 and 12 years. Income was inversely distributed across gender ($p<0.001$), with higher proportions of high income among men vs higher proportions of lower income among women. A large

proportion of women worked in routine-non manual occupations (45.5%), whereas the majority of men worked in manual occupations (55.4%) ($p<0.001$) (table 1).

According to the assessments made by the IME physicians, 26% of the men and 21% of the women received insufficient follow-up by their regular GP (table 1). Sick leave due to MSD accounted for about half of the diagnoses (46%), followed by 'other' and 'mental' diagnoses (27.7% and 25%, respectively) (table 1). Older age was significantly associated with sufficient follow-up by the regular GP, whereas country of birth and work in routine non-manual and manual occupations were significantly associated with insufficient follow-up compared with the reference groups (table 2).

We selected age, country of birth and occupational class as potential confounders in the adjusted GLM. The effective sample in the adjusted GLM was $n=846$. According to the assessments of the IME physicians, men with mental sick leave diagnoses received insufficient follow-up by the GP compared with the reference group (table 3).

This result was robust for adjustments (RR=1.8, 95% CI=1.15–2.68). Missing among variables were: income (1%), sick leave diagnoses (1.2%), occupation (5.7%) and follow-up actions (3.2%).

DISCUSSION

Data from the Norwegian IME trial and the IME physicians' assessments were used to examine the sufficiency of GPs follow-up actions, across intersectional patient groups defined by gender and major sick leave diagnoses, within 6 months of sick leave. We found strong agreement between the GPs' follow-up actions and the IME physicians' assessment of the sufficiency of these actions. However, the IME physicians assessed the GPs' follow-up of men with mental sick leave diagnoses as *not sufficient*. This finding was independent of patients' age, country of birth and occupational class.

Regardless of gender, individuals with mental sick leave diagnoses expect that GPs listen to their stories according to a British study of help-seeking behaviour and access to primary care.³¹ However, strict administration of time per consultation among GPs may act as a barrier for presenting sensitive stories about mental issues.³¹ Following this line of argument, shortage of time in the consultation may limit the GPs possibility to understand the patients' challenges and discuss appropriate follow-up actions. The important question in the current study is why follow-up actions among men with mental sick leave diagnoses alone is assessed as insufficient by the IME physicians. Taking a gender perspective, women, in general, have a higher lifetime healthcare use than men.³² Findings across 10 European countries confirm this trend.³² One reason for our findings may be that women have communicated their problems to healthcare personnel and thus received sufficient follow-up from their regular GPs. In contrast, men have less experience with communicating

Table 1 Characteristics of the study population and differences between men and women

	Total	Men	Women	Gender difference
	n (%)	n (%)	n (%)	P value
Gender	937	399 (42.6)	538 (57.4)	
Marital status				
Married (or not)	466 (49.79)	204 (51.1)	262 (48.7)	0.462*
Single (or not)	315 (33.6)	139 (34.8)	176 (32.7)	0.496*
Divorced (or not)	156 (16.6)	56 (14.0)	100 (18.6)	0.064*
Country of birth				<0.001*
Norway	793 (84.6)	301 (75.4)	456 (84.8)	
Western	76 (8.1)	52 (68.4)	24 (31.6)	
Non-Western	68 (7.3)	32 (47.1)	36 (53.9)	
Diagnosis				0.066*
L (musculoskeletal)	432 (46.1)	200 (50.1)	232 (43.1)	
P (mental)	235 (25.1)	94 (23.6)	141 (26.2)	
other	259 (27.6)	98 (24.6)	161 (30.1)	
Income last year (percentiles)				<0.001*
NOK≥517 644	232 (25)	142 (35.9)	90 (16.9)	
NOK 422 448–517 643	232 (25)	109 (27.6)	123 (23.1)	
NOK 329 402–422 447	232 (25)	84 (21.3)	148 (27.8)	
NOK≤329 401	232 (25)	60 (15.2)	172 (32.3)	
Occupational class				<0.001*
Managers/professionals	146 (15.6)	65 (16.3)	81 (15.1)	
Associate professionals	197 (21.0)	44 (11.0)	153 (28.4)	
Routine non-manual	304 (32.4)	60 (15.0)	244 (45.4)	
Manual	276 (29.5)	221 (55.4)	55 (19.2)	
(In)sufficient follow-up, assessed by IMEs				0.097*
Yes	692 (73.9)	248 (71.2)	408 (75.8)	
No	215 (22.9)	102 (25.6)	113 (21.0)	
Groups defined by gender and diagnoses				
Women/musculoskeletal	221 (24.7)			0.021
Men/musculoskeletal	191 (21.3)			
Women/mental	139 (15.5)			
Men/mental	93 (10.4)			
Women/other	157 (17.5)			
Men/other	95 (10.6)			
	Mean (SD)	Mean (SD)	Mean (SD)	
Age	46.7 (11.04)	47.0 (11.51)	46.5 (10.69)	0.520†
Sick leave 12 months pre baseline				
Days	31.5 (50.05)	28.9 (51.10)	33.45 (49.21)	0.167†

P<0.05 when numbers in bold.

* χ^2 .

†t-Test.

IME, independent medical evaluation.

their mental health problems³³ thus leaving the GPs with insufficient information to provide sufficient follow-up.

Moreover, given that men are less experienced in seeking help and more reluctant to acknowledge barriers for help-seeking,³³ they may present with more severe

illnesses when finally consulting the GP which may limit the GPs follow-up alternatives. In a large cross-sectional study, Vesga-López *et al*³⁴ found considerable gender difference in comorbidities to anxiety. Generally, men with a generalised anxiety disorder had significantly

Table 2 Univariable associations between population characteristics and the IME physicians' assessment of the (in)sufficiency of follow-up by the regular GP

	Insufficient follow-up (sufficient=reference)	
	RR	95% CI
Gender		
Women	1	
Men	1.22	0.97 to 1.54
Age		
Marital status		
Married (or not)	0.90	0.66 to 1.22
Single (or not)	1.24	0.90 to 1.70
Divorced (or not)	0.85	0.56 to 1.30
Country of birth		
Norway	1	
Western	1.17	0.78 to 1.75
Non-Western	1.47	1.02 to 2.12
Diagnosis		
L (musculoskeletal)	1	
P (mental)	1.26	0.87 to 1.83
Other	0.94	0.65 to 1.38
Income last year (percentiles)		
NOK \geq 517 644	1	
NOK 422 448 to 517 643	1.07	0.70 to 1.62
NOK 329 402 to 422 447	0.72	0.46 to 1.13
NOK \leq 329 401	0.77	0.49 to 1.19
Occupational class		
Managers/professionals	1	
Associate professionals	1.38	0.78 to 2.42
Routine non-manual	1.84	1.10 to 3.08
Manual	1.76	1.04 to 2.97
Sick leave 12 months pre baseline		
Days	1.00	1.00 to 1.00

Generalised linear models with relative risks (RRs) and 95% CIs. P<0.05 when numbers in bold. IME, independent medical evaluation.

higher rates of comorbid alcohol and drug use disorders, nicotine dependence, and antisocial personality disorder, whereas women had higher rates of comorbid mood disorders (except bipolar disorder) and anxiety disorders (except social anxiety disorder).³⁴ Possibly, there are less follow-up actions suitable for workers with comorbid alcohol and drug use disorders than for comorbid mood and anxiety disorders. It is worth noting that low level of healthcare use among Australian men with mental and substance use disorder have prompted policy initiatives to reduce stigma and increase help-seeking.³⁵ Although

these findings relate to help-seeking, they direct attention to the stigma related to comorbid alcohol and drug use disorders that may act as a barrier for open communication with the GP and thereby the relevance of follow-up actions initiated. The IME physicians, however, may have managed to detect the insufficiency of follow-up actions, due to prolonged consultation time (1 hour) and updated information on available follow-up actions in the social security system.

Another possible explanation for the IME assessed insufficiency of follow-up actions provided by the GP is that these patients may have rejected help for their mental health problems. The social consequences of having a mental health problem have been found to be worse for men due to cultural norms and expectations of 'being strong', 'active' and 'less emotional'.³⁶ A review by Samulowitz *et al*³⁷ found that men with chronic pain recognised their diagnosis as a 'women's disease' and questioned it, ignored it, or did not talk about it. This also resulted in low compliance with physicians' advice, possibly, as it jeopardised their masculinity. Thus, masculinity norms that negatively influence acceptance of need for help among patients with mental sick leave diagnoses may explain why the IME physicians considered GP follow-up actions as insufficient.

In sum, we suggest that factors related to the GPs consultation practice interact with attitudes to health seeking among men with mental sick leave diagnoses, and generate the higher risk of not receiving sufficient follow-up actions by the GP. IME physicians may discover this negative interaction due to their generous consultation time and thorough education in available follow-up actions.

Finally, it is important to bear in mind that non-significant findings among the remaining exposure groups (compared with the reference), not necessarily means that they received sufficient follow-up by their GP. It is the substantially poorer follow-up of men with mental sick leave diagnoses that shapes the pattern of significance.

Strengths and limitations

Patients allocated to the NIME-trial were drawn from the national sick leave register by NAV and included a 1 year county cohort of all workers sick-listed for the past 6 months. Moreover, the IME reports enclose all 937 patients attending the IME intervention. The random allocation ensured high generalizability to long-term sick-listed workers in Norway. However, due to the cancellation of appointments (26% of the allocated sample) and 16% not showing up, the external validity may be weakened.²²

A strength is that the reports from the IME physicians are trustworthy. Physicians engaged in the NIME-trial were experienced GPs with specialised training in different sick leave measures and RTW actions.³⁸ One important quality aspect was the 2-hour timeframe of the whole IME consultation. Furthermore, the IME physicians had in-depth access to the patients' case history

Table 3 Associations between groups defined by gender and sick leave diagnoses, and the IME physicians' assessment of the (in)sufficiency of follow-up initiated by the regular GP past 6 months

	Crude		Adjusted	
	RR	95% CI	RR	95% CI
Women/musculoskeletal diagnoses	1		1	
Men/ musculoskeletal diagnoses	1.4	0.96 to 1.96	1.4	0.92 to 2.09
Women/mental diagnoses	1.1	0.73 to 1.68	1.2	0.75 to 1.77
Men/mental diagnoses	1.8	1.24 to 2.68	1.8	1.15 to 2.68
Women/other diagnoses	1.3	0.87 to 1.87	1.3	0.88 to 1.92
Men/other diagnoses	0.9	0.51 to 1.46	1.0	0.58 to 1.73

Generalised linear models with relative risks (RRs) and 95% CIs, crude and adjusted for age, country of birth and occupational class. P<0.05 when numbers in bold. IME, independent medical evaluation.

from the regular GP and/or from the local NAV record as basis for their assessment.

A final strength is that background data on diagnoses, previous sick leave, occupation, and income were drawn from NAV registers containing little missing data. Register data ensure access to the actual sick leave diagnosis; one might, however, question how reliable this diagnosis is.³⁹ The complexity of long-term disability is associated with both medical and non-medical factors,⁴⁰ and many have comorbid illnesses. Musculoskeletal complaints often coexist with mental health problems.^{41 42} However, due to the shame and stigma associated with mental diseases,⁴³ patients may argue for a sick leave diagnosis related to musculoskeletal symptoms and not mental symptoms.

A potential limitation is linked to the 918 participants that did not receive an invitation letter due to lack of capacity among the independent medical evaluators. Analyses in the main report²² find no statistically significant differences between these and the participants who received an invitation letter, with respect to gender, age and previous sick leave. We do not have information on sick leave diagnoses among those who did not receive an invitation letter. However, there is no reason to believe that the distribution of diagnoses should be substantially different among these compared with those receiving the invitation letter. It is worth noting that the randomisation was successful, meaning that there was no statistically significant difference in relevant characteristics (gender, age and sick leave diagnoses) among the intervention group vs the control group.²²

Moreover, non-participation among those invited to the study (n=761) may cause selection bias. Leaning on results from the main report³⁷ non-participants were on average 1.5 years younger than the participants, whereas the distribution across gender and previous sick leave did not differ. Unfortunately, we do not have information on sick leave diagnoses among non-participants. However, analyses of non-participation in population-based studies indicate an association between male gender, older age, lower education and mental diagnoses.^{44 45} If this is the case in the current study, there might be an overweight

of mental diagnoses among the male non-participants. Since non-participants often have poorer health than participants and may be marginalised both in healthcare as well as work-life, there is a chance that they are not followed-up any better than those participating in the current study. If this is the case, it is likely that we have underestimated the risk among men with mental diagnoses. Finally, participation bias is most likely a greater threat to the validity of prevalence studies than to studies of associations.⁴⁶

CONCLUSION

According to IMEs in Norway, men, sick-listed for 6 months with a mental diagnoses did not receive sufficient RTW follow-up by the GP. Knowledge on vulnerable groups can help GPs to improve their practice for patients with uncovered needs. The regular GP may also benefit from receiving the same training and in-depth knowledge of different follow-up actions, aiming at RTW, as the IME physicians in the current study. Efforts should be made to clarify the needs of these men and initiate relevant follow-up actions, preventing marginalisation in working life.

Acknowledgements The authors thank Olav Hahn at the Norwegian National Advisory Unit on Occupational Rehabilitation, Rauland, Norway for valuable help with organising data from the independent medical evaluation report linking them to the national register of sickness benefits from Norwegian Labour and Welfare Administration.

Contributors All authors contributed to the design of the study and interpreted the data. IH and IØ made the statistical analyses. All authors wrote the first version of the manuscript and revised the manuscript for important content. All authors have read and approved the final version of the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval The Regional Committee for Medical and Health Research Ethics (REK) concluded that the NIME trial was not regulated by the Health Research Act

(2015/506), since the main aim of the NIME trial was to examine if the intervention reduced sickness absence (ie, an insurance outcome). The mandate of REK is strictly governed by the Health Research Law and they only evaluate projects that the law define as health research. The primary outcome measure in the NIME trial was return to full or partial work and therefore not deemed health related by the committee. The Insurance Law Act of 1997-02-28-19-§25-13 enabled the trial to be performed.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available. Data are derived from Norwegian National registries after an application process. Sharing is not allowed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Irene Øyeflaten <http://orcid.org/0000-0003-4356-5469>

REFERENCES

- Wynne-Jones G, Mallen CD, Main CJ, *et al*. What do GPs feel about sickness certification? A systematic search and narrative review. *Scand J Prim Health Care* 2010;28:67–75.
- Aamland A, Maeland S. Primary care physicians' attitudes to and experiences with sick leave and follow-up a scoping review. *Tidsskr Velferdsvetenskap* 2016;19:148–78.
- Hulshof C, Pransky G. The Role and Influence of Care Providers on Work Disability. In: Loisel P, Anema JR, eds. *Handbook of work disability*. New York: Springer, 2013: 203–15.
- Brage S, Kann I. *Fastlegers sykmeldingspraksis I: Variasjoner (GP's sickness certification practice: variations)*. Oslo: Rikstryktingen, 2006.
- Husabo E, Monstad K, Holmås TH, *et al*. Protocol for the effect evaluation of independent medical evaluation after six months sick leave: a randomized controlled trial of independent medical evaluation versus treatment as usual in Norway. *BMC Public Health* 2017;17:573.
- Koopmans PC, Roelen CA, Bültmann U, *et al*. Gender and age differences in the recurrence of sickness absence due to common mental disorders: a longitudinal study. *BMC Public Health* 2010;10:426.
- Leijon M, Hensing G, Alexanderson K. Sickness absence due to musculoskeletal diagnoses: association with occupational gender segregation. *Scand J Public Health* 2004;32:94–101.
- Mastekaasa A. The gender gap in sickness absence: long-term trends in eight European countries. *Eur J Public Health* 2014;24:656–62.
- Mastekaasa A, Melsom AM. Occupational segregation and gender differences in sickness absence: evidence from 17 European countries. *Eur Sociol Rev* 2014;30:582–94.
- Borkhoff CM, Hawker GA, Kreder HJ, *et al*. The effect of patients' sex on physicians' recommendations for total knee arthroplasty. *CMAJ* 2008;178:681–7.
- Bösner S, Haasenritter J, Hani MA, *et al*. Gender bias revisited: new insights on the differential management of chest pain. *BMC Fam Pract* 2011;12:45.
- Bugiardini R, Estrada JLN, Nikus K, *et al*. Gender bias in acute coronary syndromes. *Curr Vasc Pharmacol* 2010;8:276–84.
- Chang AM, Mumma B, Sease KL, *et al*. Gender bias in cardiovascular testing persists after adjustment for presenting characteristics and cardiac risk. *Acad Emerg Med* 2007;14:599–605.
- Crilly M, Bundred P, Hu X, *et al*. Gender differences in the clinical management of patients with angina pectoris: a cross-sectional survey in primary care. *BMC Health Serv Res* 2007;7:142.
- Hamberg K, Risberg G, Johansson EE, *et al*. Gender bias in physicians' management of neck pain: a study of the answers in a Swedish national examination. *J Womens Health Gend Based Med* 2002;11:653–66.
- Nyberg F, Osika I, Evengård B. "The Laundry Bag Project"--unequal distribution of dermatological healthcare resources for male and female psoriatic patients in Sweden. *Int J Dermatol* 2008;47:144–9.
- OECD. Sickness, Disability and Work: Breaking the Barriers. In: *A synthesis of findings across OECD countries*. OECD Publishing, 2010.
- Album D, Westin S. Do diseases have a prestige hierarchy? A survey among physicians and medical students. *Soc Sci Med* 2008;66:182–8.
- Wijnhoven HAH, de Vet HCW, Picavet HSJ. Prevalence of musculoskeletal disorders is systematically higher in women than in men. *Clin J Pain* 2006;22:717–24.
- The Norwegian Ministry of Health and Care Services. *The introduction of a regular GP scheme*. Oslo, 2001.
- Aamland A, Husabo E, Maeland S. Independent medical evaluation for sick-listed patients: a focus group study of GPs' expectations and experiences. *BMC Health Serv Res* 2018;18:666.
- Silje M. *Forsøk Med NY medisinsk vurdering (NMV) etter seks måneders sykmelding (new medical evaluation (NMV) after six months sickness absence)*. Bergen, Norway: Norce - Norwegian Research Centre, 2018.
- World Health Organisation. *International classification of primary care (ICPC-2)*. 2 edn, 2003.
- Norway S. *International Standard Classification of Occupation (ISCO-88(COM))*, 1999.
- Eriksson R, Concepts GJH. Data, and Strategies of Enquiry. In: *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford: University of Oxford, 1992: 28–64.
- Krokstad S, Westin S. Disability in society-medical and non-medical determinants for disability pension in a Norwegian total County population study. *Soc Sci Med* 2004;58:1837–48.
- Leinonen T, Pietiläinen O, Laaksonen M, *et al*. Occupational social class and disability retirement among municipal employees--the contribution of health behaviors and working conditions. *Scand J Work Environ Health* 2011;37:464–72.
- Grant RL. Converting an odds ratio to a range of plausible relative risks for better communication of research findings. *BMJ* 2014;348:f7450.
- Talbot D, Massamba VK. A descriptive review of variable selection methods in four epidemiologic journals: there is still room for improvement. *Eur J Epidemiol* 2019;34:725–30.
- Hernán MA, Hernández-Díaz S, Werler MM, *et al*. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. *Am J Epidemiol* 2002;155:176–84.
- Bristow K, Edwards S, Funnell E, *et al*. Help Seeking and Access to Primary Care for People from "Hard-to-Reach" Groups with Common Mental Health Problems. *Int J Family Med* 2011;2011:1–10.
- Kovess-Masfety V, Boyd A, van de Velde S, *et al*. Are there gender differences in service use for mental disorders across countries in the European Union? results from the EU-World mental health survey. *J Epidemiol Community Health* 2014;68:649–56.
- Galdas PM, Cheater F, Marshall P. Men and health help-seeking behaviour: literature review. *J Adv Nurs* 2005;49:616–23.
- Vesga-López O, Schneier FR, Wang S, *et al*. Gender differences in generalized anxiety disorder: results from the National epidemiologic survey on alcohol and related conditions (NESARC). *J Clin Psychiatry* 2008;69:1606–16.
- Harris MG, Diminic S, Reavley N, *et al*. Males' mental health disadvantage: an estimation of gender-specific changes in service utilisation for mental and substance use disorders in Australia. *Aust N Z J Psychiatry* 2015;49:821–32.
- Hensing G, Alexanderson K, Allebeck P, *et al*. Sick-leave due to psychiatric disorder: higher incidence among women and longer duration for men. *Br J Psychiatry* 1996;169:740–6.
- Samulowitz A, Gremyr I, Eriksson E, *et al*. "Brave Men" and "Emotional Women": A Theory-Guided Literature Review on Gender Bias in Health Care and Gendered Norms towards Patients with Chronic Pain. *Pain Res Manag* 2018;2018:1–14.
- Aamland A, Øyeflaten I, Maeland S. Independent medical evaluation for sick-listed workers in Norway: a focus group study of the experience of IME doctors. *Scand J Public Health* 2019;47:70–7.
- Maeland S, Werner EL, Rosendal M, *et al*. Sick-leave decisions for patients with severe subjective health complaints presenting in primary care: a cross-sectional study in Norway, Sweden, and Denmark. *Scand J Prim Health Care* 2013;31:227–34.
- Nilsen SM, Ernsten L, Krokstad S, *et al*. Educational inequalities in disability pensioning - the impact of illness and occupational, psychosocial, and behavioural factors: The Nord-Trøndelag Health Study (HUNT). *Scand J Public Health* 2012;40:133–41.
- Ihlebaek C, Brage S, Eriksen HR. Health complaints and sickness absence in Norway, 1996-2003. In: *Occupational medicine*. England: Oxford, 2007: 57. 43–9.
- Reme SE, Tangen T, Moe T, *et al*. Prevalence of psychiatric disorders in sick listed chronic low back pain patients. *Eur J Pain* 2011;15:1075–80.
- Knapstad M, Øverland S, Henderson M, *et al*. Shame among long-term sickness absentees: correlates and impact on subsequent sickness absence. *Scand J Public Health* 2014;42:96–103.
- Knudsen AK, Hotopf M, Skogen JC, *et al*. The health status of nonparticipants in a population-based health study: the Hordaland health study. *Am J Epidemiol* 2010;172:1306–14.



45 Korkeila K, Suominen S, Ahvenainen J, *et al.* Non-Response and related factors in a nation-wide health survey. *Eur J Epidemiol* 2001;17:991–9.

46 Knudsen AK, Øverland S, Hotopf M, *et al.* Lost working years due to mental disorders: an analysis of the Norwegian disability pension registry. *PLoS One* 2012;7:e42567.