Does variation in GP practice matter for the length of sick leave? A multilevel analysis based on Norwegian GP-patient data

Abstract

In Norway, as in many countries, the national insurance system is under economic stress from demographic change impacting on the pensions versus contributions balance, and an increasing number of disability and sickness benefit claimants. The general practitioner (GP) is responsible for assessing work capacity and issuing certificates for sick leave based on an evaluation of the patient. Although many studies have analyzed certified sickness absence and predictive factors, no studies assess its variation between patients, GPs or geographical areas within a multilevel framework. Using a rich Norwegian matched patient-GP data set and employing a multilevel random intercept model, the study attempts to disentangle patient, GP and municipality-level variation in the certified sickness absence length for Norwegian workers in 2003. We find that most observed patient and GP characteristics are significantly associated with the length of sick leave (LSL) and medical diagnosis is an important observed factor explaining certified sickness durations. However, 98% of the unexplained variation in the LSL is attributed to patient factors rather than influenced by variation in GP practice or differences in municipality-level characteristics. Our findings indicate that GPs practice variation does not matter much for the patients' LSL. Our results are compatible with a high degree of patient involvement in current general practice. Based on this understanding one may infer that GPs play an advocate role for their patients in Norway, where the patients' own wishes are important when decisions are made.

Keywords: Norway; general practitioners (GPs); length of sick leave; multilevel regression models; matched GP-patient data

Introduction

In Norway, as in many countries, the national insurance system is under economic stress from demographic change impacting on the pensions versus contributions balance, and an increasing number of disability and sickness benefit claimants. At a given working day around 6.5% (130,000 persons) of the workforce received sickness benefits from the National Insurance Scheme in 2008 (Nav, 2008). The number of persons on sickness benefits is high in Norway compared with other countries (Bonato and Lusinyan, 2004).

Three institutional factors may partly explain the high number of persons on sickness benefits. First, Norway has generous sickness benefits, paying 100% of the current wage for up to a year. Second, the cost to employers of having workers on sick leave is low. Third, the general practitioners' (GPs) medical assessments are seldom scrutinized or evaluated by social insurance institutions.

The Norwegian sickness system provides both cash benefits and medical benefits within the social insurance system. Employers pay cash benefits for the first 16 calendar days of sick leave, while the insurance system covers the wage loss from the 17th day up to a maximum of 52 weeks. After that period the patient is covered by a rehabilitation allowance or a disability pension. Self-certification can only be used within the first three days per spell (a total of 12 days in a 12 month period), and the employee must obtain a medical sickness certificate from a physician from the fourth day of the sickness spell. If a person works in an IA firm (firms with a special agreement with the government, in which one aim is to reduce sick leave), the self-certification is eight calendar days per spell (a total of 24 days in a 12 month period). For longer sick leaves, the employee and the employer must work out a rehabilitation plan within six weeks.

Norwegian general practice is a list-based system in which every inhabitant has the right to be in the care of a GP. GPs are allowed to have up to 2500 patients on their lists but may limit their lists below this level. The payment system (salary model) is a mix of a capitation fee and fee-for-service. The GP has a duty to prioritize inhabitants on his/her own patient list, and the GP scheme formalizes the relationship between the patient and doctor. GPs have the responsibility for planning and coordinating individualized preventive work, examination and treatment. They are also responsible for the patient's medical records, for updating medical histories and recording medicine use.

Previous studies identify that main predictive factors for the length of sick leave (LSL) are health (diagnosis), age, gender, family circumstances, economic incentives and restrictions in insurance legislation, type of work, social norms and the functioning of the labor market; see for instance, Tellnes (1989), Alexanderson (1998), Shiels and Gabbay (2006). In a review of the literature, Söderberg and Alexanderson (2003) found that, given the same patient characteristics, there are large differences between GPs regarding the LSL they certified. In addition, there is a large degree of variation in different geographical regions such as municipalities (Mabeck and Kragstrup, 1993; Arrelöv *et al.,* 2005).

Medical diagnoses related to sickness absence are seldom available; see Feeney et al. (1998). The validity of the diagnoses can vary both because GPs do not always check what is the correct diagnosis or they want to protect the patient. The International Classification of Primary Care (ICPC) has been used to diagnose sickness absence episodes in Norway since 1992. In Norway, most sickness certificates are issued by GPs (Kristensen, 1993; Brage et al., 1996). Studies have shown that GPs are largely homogenous in terms of sickness certification especially for main diagnostic groups (Hofmans-Okkes and Lamberts, 1996; Britt, Angelis, and Harris, 1998).

GPs act as gatekeepers in the sickness benefit system in most western European countries, see for instance, Stone (1979), Meershoek *et al.* (2007), Swartling *et al*. (2007). GP responsibilities include assessing ability to work and issuing certificates for sick leave based on an evaluation of the patient. This means that they assess the existence of disease, decide whether the disease affects the ability to work, and weigh the pros and cons of sick leave. If sick leave is recommended, the GP must decide its duration and grade (full or part time), and measures to be taken during the absence such as treatment, rehabilitation, medication, contact with the employer, referrals and examinations (Söderberg and Alexanderson, 2003). The GP must then do paperwork, such as issuing a certificate stating the medical diagnosis and the duration, and recommending activities or rehabilitation measures. Sometimes more information is required, which makes the evaluation of the patient and her situation complex and time consuming.

Though many studies have analyzed certified sick leave and predictive factors, very few studies focus on the LSL within a multilevel statistical model where variation in sickness absence length are analyzed at different levels (i.e. patient/GP/municipality) (Söderberg and Alexanderson, 2003). To date, very little attention has been given to understanding the gatekeeping role of GPs and how much of the variation in sickness absence they can explain. Some degree of patient involvement is the rule in current general practice. To shed light on the issue using a rich administrative matched patient-GP data set, the paper aims to analyze factors explaining the variation in the length of certified sick leave and to disentangle patients, GPs and municipality sources of explained and unexplained (residual) variation in the certified sickness durations for the population of Norwegian workers.

Methods

Data and variables

This study used a rich Norwegian administrative data set in the analysis of how individual LSL is affected by the individuals' personal characteristics, GP characteristics and municipality attributes. In the study, we merged data from the following data sources. First, information on individual sick leave, together with extensive individual background information, was taken from the FD-Trygd database in Statistics Norway (SSB). This database contains social insurance information on the entire Norwegian population. Second, data on GPs (e.g. age, gender, and patient list length) were collected from the Norwegian labor and welfare organization's (NAV's) regular GP database, where each patient is connected to one GP. Thus, by merging the two data sets we can explore extensive information on both GPs and their patients. Third, information on municipality-specific characteristics is obtained from Statistics Norway (SSB). Table 1 provides a description of all variables used in the analysis.

The study uses pseudo-anonymous patients' and GPs' information and the research project of which the study is a part, has been approved by the Regional Committee for Medical and Health Research Ethics (REK).

The sample

From the FD-Trygd database we have extracted information on all employees (423,022) who started a certified sick leave episode in 2003 compensated by the National Insurance Administration (NIA). Episodes ending in 2004 are included as long as they start in 2003. After dropping individuals with missing information, our sample consists of 291,971 individuals and 353,611 sickness episodes. 243,896 individuals had one episode of sickness absence, 40,239 had two episodes, 6,041 had three episodes, while 1,795 individuals had more than three episodes. Each episode is treated as an independent occurrence in our empirical analysis.

Dependent variable

The length of sickness absence (LSL)

The FD-Trygd database contains all sickness absence compensated by the National Insurance Administration (NIA). In Norway, the first 16 working days of a sickness episode are compensated by the employer and sickness absence shorter than 17 days are not included in the database. The maximum duration of a sick leave episode is one year and the outcome measure in this study is the length of certified sickness absence (LSL) between 17 days and one year.

Independent variables

As seen in Table 1, as predictors of the LSL we have considered three types of observable attributes - patient, physician and municipality.

Patient characteristics

As patients observable attributes we consider patient socio-demographic characteristics (i.e. age, gender, income, education, household size and number of children in the household). Patient work-related characteristics, namely working hours per week, experience and the industry where the patient worked are also used as control variables. Moreover, as patient-level covariates we control for diagnosis (health) using dummy variables for 16 main chapters in ICPC (sickness absence due to social problems - chapter Z in ICPC - is not permitted to use in sickness certificates in Norway). We also include the total number of sickness days in 2002 (for those days covered by the NIA) as a control variable, since this information may be used as a proxy of past health status.

Physician characteristics

Physician-level attributes include age, gender and whether or not the GP is a specialist (in general medicine or community medicine). To control for the workload of a GP, we include the GP's list length in our model. GPs are either paid a fixed wage or through a capitation scheme. This could influence the LSLs and we include a dummy variable for this in our regression model (Aakvik and Holmås, 2006).

Municipality characteristics

To control for observable area-level characteristics, we use municipality-level attributes as covariates. To control for municipalities' location we introduce three dummy variables considering whether patients live in a municipality that are located in large cities (e.g. Oslo, Bergen), or in urban areas or rural areas. The dummy variable for 'large cities' is used as a reference category in the regression model. We further include two municipality-level attributes - index mortality and index unemployment (see Table 1) as control variables.

Estimation strategy

Our statistical analysis anticipates that the patients' LSL is partly dependent on patients' attributes, physician (GP) characteristics and the administrative municipalities to which they belong. This hierarchical or nested structure in patients' LSL is modeled by separating the patient, GP and municipality sources of variation. To account for the clustering/nested effects at higher levels (i.e. individuals nested within GP and municipalities) in our modeling approach, we assume that the coefficients for all three levels are fixed but that the intercept is randomly varying (i.e. assuming a random-intercept model) (Cameron and Trivedi, 2005; Rabe-Hesketh and Skrondal, 2005). We may write the following multilevel/nested linear model:

$$y_{igm} = \beta X_{igm} + \gamma Y_{gm} + \delta Z_m + w_m + u_{ig} + e_{igm} \,,$$

where y_{igm} represents the LSL (the dependent variable), which is related to a vector of patient-level explanatory variables *X*, a vector of GP characteristics *Y*, and a vector of municipality characteristics Z.

In this specification the overall error term v_{igm} is decomposed into $w_m + u_{ig} + e_{igm}$, where w_m is the random error term for the *m*th municipality, u_{ig} denotes the nested effect of *i*th individual within the *g*th GP and e_{igm} is the remaining disturbance term (error term for i*th* patient treated by the g*th* GP within the *m*th municipality). The error terms are assumed to be iid with mean zero.

The ability to partition variance at different levels (e.g., municipality, GP and patient) is a unique feature of multilevel regression analysis. The model is useful for estimation and quantification of the relative importance of individual and higher-level effects (e.g. in our case patients nested within GPs) and for understanding patient/GP/municipality variations in the length of sick leave (Snijders and Bosker, 1999). In our 3-level multilevel regression analysis, where Level-1 contains 291,971 patients (in total 353,611 episodes) nested within 3,709 GP at level-2, and GPs are nested within 415 municipalities at level-3. Because the dependent variable is continuous, a multilevel linear model is used.

To examine the variations in LSL explained by different levels, ideally four sequential models could be estimated. The first considers a null (empty) model of patients (level 1) nested within GP (level 2) and GP nested within Municipalities (level 3) with no variables in the fixed part and only the intercepts in the random part of the model (Model I). This model is used as a baseline for comparing the size of higher levels variations (e.g. GP variations) in the patients' LSL in subsequent models. In the second model we add all the patients' characteristics in the fixed part of the model (Model II). Model III is the same as model II, but also adds GP characteristics in the fixed part of the model. Controlling for patient characteristics, this model potentially examines the effect of GP-level predictors on the patients' length of sick leave (fixed part). In the random part of the model the practice variation of the GPs certification on the length of sick leave is estimated before and after taking into account the effect of the GP-level observable characteristics. Finally, the fourth model not only includes all patient and GP-level predictors but also add municipality-level observable characteristics in the fixed part of the model. In the random part of the results, this model facilitate us to examine the extent to which municipal observable characteristics explains municipality-level differences in the patients' length of sick leave (Model IV).

To illustrate the relevance of the GP or municipality differences (variances) for understanding the patients' differences in the LSL, we calculate the intra class-correlation (ICC). The intra-class (cluster) correlation can be expressed as the proportion of the patient differences in the LSL (i.e., patient-level variance) that is at the GP or municipality-level. For example, the proportion of the patient-level variance ($\sigma_M^2 + \sigma_G^2 + \sigma_i^2$) that is at the GP-level (σ_G^2) can be calculated by the general formula as $ICC = \sigma_G^2 / (\sigma_M^2 + \sigma_G^2 + \sigma_i^2)$. The closer the ICC is to 0%, the smaller proportion of the total variance is at the GP or the municipality-level, implying a low relevance of the GP or municipality factors for understanding patient disparities in the LSL. By this strategy, in particular, we are able to quantify how much of the GP differences in the LSLs are explained by differences in the patients composition of the GP or municipality, and how much of these GP or municipality differences are explained by the GP characteristics or the municipal-level of attributes.

Results

Descriptive analysis

Table 2 provides descriptive statistics for the variables used in the statistical models. We first notice that the data consist of considerably more females (215,087) than males (138,524). The average age is 42 years for both the genders. We observe that the average LSL is just over 62 days. There are only small gender differences; males have about two days longer sick leave than females. Many of the patients in our sample have not been on sick leave in 2002; therefore, the average total number of sick days in 2002 is around 19 days for our group of patients.

For all sickness episodes, more than 60% of sickness absence is within the diagnosis groups 'Musculoskeletal' and 'Psychological' problems (two main chapters in ICPC). Notice that nearly half of the males' certified sickness absence is reported due to musculoskeletal related health problems, whereas, for the same broad diagnosis group, the proportion is 40% for the females. Psychological problems are more prevalent among females than males (17% versus 15%), whereas for other broad diagnoses groups (i.e. main chapters in ICPC) no noteworthy gender difference is observed (see Table 2).

Regarding GP characteristics, as seen in Table 2, the average age of the GPs is around 48 years and 72% of patients' sickness leave is certified by the male GPs. The average list length of the GPs is 1335 and around 6% of the GPs are on fixed salary. Concerning municipality attributes, roughly 58% of the patients live in municipalities located in large cities (e.g. Oslo and Bergen), 29% in other urban areas and the rest live in municipalities situated in rural areas. Average index mortality and index unemployment are found to be 5.6 and 6.1, respectively.

Table 3 illustrates the average LSL per 16 main chapters in ICPC. In 10 out of the 16 main chapters, the average LSL are higher for the males than females. The highest average LSL is observed for the 'Blood' related diagnosis (86 days) and the lowest LSL is found under 'Respiratory' diagnosis (29 days). In both cases the LSLs are higher for the males than females. For the two main chapters in ICPC (i.e. 'Musculoskeletal' and 'Psychological'), the average LSL is around 66 and 76 days, respectively. There are only minor gender differences in the LSL for these diagnoses. In connection to the 'Musculoskeletal' diagnosis group the average LSL is higher for females (69 days versus 64 days), whereas for the 'Psychological' group the average LSL is higher for the males than females (79 days versus 74 days).

Analytical results

We first present the fixed-part results followed by the random-part results of the random intercept model. Note that, because of minor changes in the fixed part results based on models II and III, we do not present the results for these models in Table 4. Instead, we discuss the model with all three level predictors (Model IV) in the fixed part of the results. In addition to presenting the full model, we also report results for male and female patients separately.

Fixed-part results

Patient-level characteristics

The effects of the patient-level covariates on the LSL are presented in Table 4, which shows that most of the patient-level predictors are significantly associated with the LSL at the 1% level. In particular, the results show that age is positively and significantly associated with the LSL for both genders. The effect of age is higher for males than for females. As expected, different socioeconomic variables, such as education, income, work experience, and working hours, are negatively and significantly associated with the LSL, and the effects are stronger for males than females.

Comparing patients using ‘General and unspecified (A)’ ICPC as the base category, we find that the LSL is significantly higher for 10 of the ICPC chapters, and significantly shorter for 5 of the ICPC categories (i.e. ‘Digestive (D)’, ‘Eye (E)’, ‘Respiratory (R)’, ‘Skin (S)’ and ‘Urology’), see Table 4 for more details. Relative to a patient working in manufacturing industry, LSLs are significantly higher for patients working in agriculture, construction, wholesale and retail or financial sectors.

Physician-level attributes

As seen in Table 4, only few of the GP-level attributes are significantly associated with the patients' LSL. A fixed-salary GP with a longer list length issues shorter LSL certificates than their counterpart. It is interesting to note that male GPs certify shorter LSLs for male patients, on average more than one day shorter than for their female counterparts. No significant difference in the LSL is observed for the female patients when the sick leave certificate is issued by a male physician. Older GPs seem to issue longer (shorter) LSL for the males (females) and GPs with specialization (in general medicine or community medicine) seem to issue longer LSL than GP with no specialization for both male and female patients; however, none of these associations are found to be significantly different from zero.

Municipality-level characteristics

None of the municipality-level attributes are significantly associated with patients' LSL. In comparison with a municipality located in a large city, a female patient's LSL is significantly around three days shorter if she lives in a municipality located in rural area; however, this municipality characteristic is not significantly associated with the male patients' LSL. Municipality-level `index mortality' seems positively and significantly related with male patients' LSL but insignificantly related with the females' LSL.

Random-part results

To what extent is GP practice variation or municipality variation important for the patients' LSL? Table 5 describes the random part of the results which gives us indications of this question. The null model with no predictors/variables (Model-I) shows a significant variation in the LSL between GPs ($\sigma_G^2 = 36.58$) and municipalities ($\sigma_M^2 = 57.01$) for all patients. After controlling for patients' observable attributes (model-II), an insignificant decrease in the variation between GPs ($\sigma_G^2 = 35.4$) is observed, and the variation also decreases at the municipality and patient-level. After accounting for the patients' and GPs' characteristics (model-III), the variation between GPs and municipalities further decreases slightly. Finally, after controlling for patients', physicians' and municipality observable attributes (Table 5, model-IV), between-physicians variation reduced negligibly ($\sigma_G^2 = 35.2$) and between-municipality variation is found to be almost constant ($\sigma_M^2 = 48.2$), although the variability is still significant. Similar findings regarding changes in the variations from model to model are also observed for the males and females.

To quantify the importance of variation in GP practice on patients' LSL, intra-class (cluster) correlation (ICC) statistics can be used. In the random part of our results (in the null model), the ICC was 0.61% (Model I) for the GP level and 0.95% at the municipality level. After including patients' and GPs' socio-demographic predictors and municipality-level observable characteristics (Model IV) the ICC is found to be almost constant for the GP level (0.62%) and it is slightly decreased at the municipality level (0.85%). This result suggests that variation in the patients' LSL is mainly affected by patient individual attributes (more than 98%) rather than influenced by the GP-level residual variation or differences in the municipality-level characteristics where they live.

Discussion

Although many studies have analyzed certified sick leave and predictive factors, there is still little known about the factors explaining the variation in the length of certified sick leave and the relative contributions to the variation from patient, GP or geographical area where they live. This paper is one of the first that uses merged administrative data on patients and GPs to analyze how observed and unobserved factors influence the duration of LSL. Our multilevel linear random intercept model allows us to disentangle patient, GP and municipality sources of variation in LSL for the Norwegian population of workers in 2003.

We find that few of the observed GP characteristics (list length, wage scheme) contribute significantly to the variation in length of LSL. Patient factors contribute to a much larger extent than GP or municipality factors. Our results also show that more than 98% of the residual variation in the length of sickness absenteeism is attributed to patient factors rather than influences of random variation in GP practice or differences in municipality-level characteristics. Our results indicate that differences across patients certified in different practices are not as important as characteristics shared by the total group of patients itself, both in terms of observed and unobserved differences.

Previous studies reported that GPs face many dilemmas in deciding the LSL (Timpka *et al.,* 1995; Englund and Svärdsudd, 2000; Hussey *et al.,* 2004), and there are sometimes conflicting interests between the patient and the GP. In particular, Englund and Svärdsudd (2000) found that even in cases where the GP would not recommend sick leave, a certificate was issued in 87% of cases, and concluded that patients appear to have a strong influence on sick leave practice. Tellnes and colleagues (1990), however, reported a large variation between doctors.

Not only due to differences in the statistical approaches used in earlier studies, but also due to differences in the data, particularly differences in the construction of the dependent variable (i.e. the sickness absenteeism variable), for instance, duration/length versus prevalence of sickness absence, short versus long term sickness absence, single episode or all episodes, specific diagnosis or all diagnosis, etc, makes it difficult to compare the findings in the literature. Having in mind these complications, nevertheless, it may be useful to compare some of our findings (fixed-part results) with other studies.

The influence of GP gender on the issuing of sickness certificates has also been discussed earlier in the literature (Söderberg and Alexanderson, 2003). The studies reported different findings and no indicative conclusion had been found on the differential practice of male and female GPs. In the regression model, we find that male GPs certify a significantly lower LSL for male patients: on average more than one day shorter than their female counterparts. However, no significant difference in LSL is found for female patients, whether the sick leave certificate is issued by a male or female GP. Shiels and Gabbay (2006) found that sickness certification of male patients by male GPs were associated with increase prevalence of sickness certification in the intermediate sickness absence period (6-28 weeks). The authors, however, found no such significant association for longer sickness absence (28 weeks or more). In a previous comparative study, Tellnes *et al.* (1990) found that older physicians issued certificates of longer durations. Although descriptive statistics confirms this result in our study (not shown here), we do not find any significant influence of GPs age on LSL in the regression model reported in Table 4.

Existing studies on GP practice variation use simple statistical analyses and there is no study that disentangles patient, GP and municipality sources of variation in the LSL in a multilevel framework. Hence, it is not possible to compare our random-part results with other studies. Nevertheless, using a much simpler model, Shiels and Gabbay (2006) find that GP effects were much smaller than anticipated, which is in line with our results.

It is also hypothesized that GP practice variation in sick leaves may be influenced by area or structural-level factors (Söderberg and Alexanderson, 2003). Using a simple statistical analysis, Arrelöv *et al.* (2005) find a large variation of the length of the sick-leave certificates and sick-leave episodes between counties and between communities in Sweden. However, using a multilevel framework, we find a negligible (less than 1%) unexplained practice variation in the length of sickness absenteeism at the municipality level. Our finding is also supported by other studies that use Scandinavian data to asses the importance of area variations for other health measures. In particular, using a multilevel method, Islam *et al.* (2006) conclude that the variation in health status is mainly affected by individual factors (more than 98%) rather than municipality characteristics in Sweden.

While we believe our analysis offers many advantages compared with many other studies, our study is not without limitations. Using a large Norwegian register database we have investigated a wide range of patient and GP attributes potentially associated with LSL. However, this study uses a cross-sectional method and hence is limited in terms of the potential to establish causal relationships. Our study used medical diagnoses based on a common classification system, ICPC. Question could be raised about the validity of GP diagnoses as one could suspect that GPs may intentionally state incorrect diagnosis from convenience or to protect the patient. However, ICPC is a well known classification system in Norway, and is used since 1992. In particular, Britt *et al.* (1998) show that GPs to a large degree use the same ICPC diagnoses when presented to identical medical cases. GPs use of medical diagnoses on sickness certificates may not always be correct. However, no clear financial incentives should induce GPs to misclassify diagnoses. Bratberg, Gjesdal, and

Mæland (2009) conclude that the diagnostic skills and integrity of GPs are acceptable in Norway, at least for musculoskeletal impairments.

The implications of sickness certificates issued by non-GPs could be another concern. Brage and Kann (2006) have studied the variation in GP practice in Norway in 2003. They find that 79% of all sick leave certificates of more than 16 days are issued by the patients own GP. We do not know how many patients are certified by another physician than their own GP. This might weaken the link between GP characteristics and the length of sick leave in our analysis. Residential mobility and the length of time that the patient is treated by the same GP is not observed in our data. Clustering of GP practices in the same municipality could also be a concern (Islam *et al.,* 2006). We do not adjust for potential omitted variables in our analyses.

Another concern could be that the low variance between GPs found in our models underestimates the true variance of the association between patients' characteristics, so the LSL is not constant across GPs. To allow the effect of patient-level covariates to vary across GPs, we could permit the slope of the patient-level variables to vary at the GP level. We have tried to estimate a random-coefficient model for different patient-level covariates (e.g. age, education and income); however, these models do not converge and we leave further exploration of this issue for future work. We have also estimated a four-level model allowing for multiple spells within individuals. The results do not change much compared to a three-level model.

Conclusions

Medical diagnosis is an important observed factor explaining certified LSL, and this feature is an intrinsic part of the GPs' role as gatekeepers. The GP is expected to act as a rationing agent on behalf of society and to adhere to national guidelines for prescriptions and referrals. At the same time the powers and rights of patients and the public have been strengthened through several organizational and legal reforms that encourage doctors to share decisions with patients. Previous studies recognized that the gatekeeping role of GPs seems to be weak compared with the role of being a patient advocate (Ford, 1998; Berg *et al.*, 2000; Hussey *et al.,* 2004) and many GPs want to relinquish their gatekeeping role (Hussey *et al.*, 2004).

We find that 98% of the residual variation in the length of sickness absenteeism is attributed to individual factors rather than influenced by variation in GP practice or differences in municipality-level characteristics. Low variation across GPs may be understood by the fact that GPs play an advocate role for their patients in Norway, where patient involvement is strong. Based on our findings, one may infer that the gatekeeping role of Norwegian GPs is probably weak compared with their advocate role, although GPs are reasonably consistent in the way they certify patients. Patients appear to have a strong influence on sick leave practice. The immediate policy implications of our findings is that to reduce sickness absenteeism in Norway focus should be on the patients' health (diagnosis) and incentives that may influence the length of sickness absence. The GPs' gatekeeping role and its implication on the length of sickness absence should be evaluated further by social insurance institutions.

References

Aakvik, A., & Holmås, T. H. (2006). Access to Primary Health Care and Health Outcomes: The Relationships between GP Characteristics and Mortality Rates. *Journal of Health Economics*, 25, 1139-1153.

Alexanderson, K. (1998). Sickness absence: A review of performed studies with focused on levels of exposures and theories utilized. *Scandinavian Journal of Social Medicine*, 26, 241-249.

Arrelöv, B. E., Borgquist, L., & Svärdsudd, K.F. (2005). Influence of local structural factors on physicians' sick-listing practice: a population-based study. *European Journal of Public Health*, 1(5), 470-474.

Berg, M., Horstman, K., Plass, S., & van Heusden, M. (2000). Guidelines, professionals and the production of objectivity: Standardisation and the professionalism of insurance medicine. *Sociology of Health & Illness*, 22(6), 765-791.

Bonato, L., & Lusinyan, L. (2004). Work Absence in Europe. IMF Working Paper 04/193. Washington DC: International Monetary Fund.

Brage S., Bentsen B. G., Bjerkedal T., Nygård J. F., & Tellnes G. (1996). ICPC as a standard classification in Norway. *Family Practice*, 13:391-396.

Brage, S., Nygard, J. F., & Tellnes, G. (1998). The gender gap in musculoskeletal-related long term sickness absence in Norway. *Scandinavian Journal of Public Health*, 26, 34-43.

Brage, S., & Kann I. C. (2006). Fastlegers sykemeldingspraksis I: Variasjoner. RTV-rapport 05/2006. Oslo, Rikstrygdeverket. (In Norwegian).

Bratberg, E., Gjesdal S., & Mæland J. G. (2009). Musculoskeletal impairments in the Norwegian working population: the prognostic role of diagnoses and socioeconomic status. A prospective study of sickness absence and transition to disability pension. *Spine*, 34(14), 1519-1525.

Britt, H., Angelis M., & Harris E. (1998). The reliability and validity of doctor-recorded morbidity data in active data collection systems. *Scandinavian Journal of Primary Health Care*, 16, 50-55.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and application*, Cambridge University Press, Cambridge.

Englund, L., & Svärdsudd, K. (2000). Sick-listing habits among general practitioners in a Swedish county. *Scandinavian Journal of Primary Health Care*, 18(2), 81-86.

Feeney, A., North, F., Head, J., Canner, R., & Marmot, M. (1998). Socioeconomic and sex differentials in reason for sickness absence from the Whitehall II Study. *Occupational and Environmental Medicine*, 55, 91-98.
Ford, F. (1998). Sickness certification: Time to scrap the med 3?" *British Journal of General Practice*, 48, 1353.

Hensing, G., Andersson, L., & Brage, S. (2006). Increase in sickness absence with psychiatric diagnosis in Norway: A general population-based epidemiologic study of age, gender and regional distribution. *BMC Medicine*, 4(19), 1-9.

Hensing G., Brage S., Nygård J. F., Sandanger I., & Tellnes G. (2000). Sickness absence with psychiatric disorders: an increased risk for marginalisation among men? *Social Psychiatry and Psychiatric Epidemiology*, 35(8), 335-340.

Hofmans-Okkes, I. M., Lamberts, H. (1996). The International Classification of Primary Care (ICPC): New applications in research and computer-based patient records in family practice. *Family Practice*, 13, 294-302.

Hussey, S., Hoddinott, P., Wilson, P., Dowell, J., & Barbour, R. (2004). Sickness certification system in the United Kingdom: Qualitative study of views of general practitioners in Scotland. *British Medical Journal*, 328, 88-91.

Islam, M. K., Merlo, J., Kawachi, I., Lindstrom, M., Burström, K., & Gerdtham, U-G. (2006). Does it really matter where you live? A panel data multilevel analysis of Swedish municipality level social capital on individual health-related quality of life. *Health Economics Policy and Law*, 1(3), 209-235.

Kristensen, F. B., Kelstrup, J., Kohlbau, C., & Lassen, L. C. (1993). Computer-based longitudinal recording of episodes of care in general practice using the International Classification of Primary Care (ICPC). Experience from one practice. Perspectives for audit and quality assessment. *Scandinavian Journal Primary Health Care Suppl*., 11(1), 53-56.

Mabeck, C. E., & Kragstrup, J. (1993). Is variation a quality in general practice? *Scandinavian Journal Primary Health Care Suppl*., 1, 32-35.

Meershoek A., Krumeich A., & Vos R. (2007). Judging without criteria? Sickness certification in Dutch disability schemes. *Sociology of Health & Illness*, 29(4), 497-514.

Nav (2008). Statistical Yearbook. NAV

Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modeling using Stata* (second edition). Stata Press. College Station, Texas.

Shiels, C., & Gabbay, M. (2006). The influence of GP and patient gender interaction on the duration of certified sickness absence. *Family Practice*, 23(2), 246-252.

Shiels, C., & Gabbay, M. (2007). Patient, clinician, and general practice factors in long-term certified sickness. *Scandinavian Journal of Public Health*, 35(3): 250-256.

Snijders, T. A. B., & Bosker R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*, Thousand Oaks, CA: SAGE Publications.

Stone, D. A. (1979). Physicians as gatekeepers: Illness certification as a rationing device. *Public Policy*, 27(2), 227-254.

Swartling, M., Peterson, S., & Wahlström, R. (2007). Views on sick-listing practice among Swedish General Practitioners - a phenomenographic study. *BMC Family Practice*, 8(44), 1-9.

Söderberg, E., & Kristina, A. (2003). Sickness certification practices of physicians: a review of the literature. *Scandinavian Journal of Public Health*, 31(6), 460-474.

Tellnes, G. (1989). Sickness certification in general practice: A review. *Family Practice*, 6, 58-65.

Tellnes, G., Sandvik L., & Moum T. (1990). Inter-doctor variation in sickness certification. *Scandinavian Journal of Primary Health Care*, 8, 45-52.

Timpka, T., Hensing, G., & Alexanderson, K. (1995). Dilemmas in sickness certification among Swedish physicians. *The European Journal of Public Health*, 5(3), 215-219.

Table 1: Variable description

Variable name	Definition
Dependent variable	
Sick days	Number of sick days covered by the National Insurance Administration (excluding spells shorter than 16 days)
Explanatory variables	
Patient characteristics	
Male	1 if patient is male, 0 otherwise.
Age	Age of the patient
Years of education	Years of completed education
Income/1000	Labor income in 2002 (in 1000 NoK)
Individuals in the household	Number of individuals in the household
Number of children < 6	Number of children younger than 6
Years of experience	Years with income above basic counting unit in pension system in 2002 (NoK 56861)
Sick days in 2002	Number of sick days in 2002
Working hours per week	
Working hours 4-19	1 if patient is working 4 – 19 hours per week, 0 otherwise
Working hours 20-29	1 if patient is working 20 – 29 hours per week, 0 otherwise
Working hours 30 +	1 if patient is working 30 or more hours per week, 0 otherwise
Chapters in ICPC:	
General and unspecified (A)	1 if diagnosis within ICPC-chapter "General and unspecified" (A), 0 otherwise
Blood (B)	1 if diagnosis within ICPC-chapter "Blood, blood forming organs, lymphatics, spleen" (B), 0 otherwise
Digestive (D)	1 if diagnosis within ICPC-chapter "Digestive" (D), 0 otherwise
Eye (F)	1 if diagnosis within ICPC-chapter "Eye" (F), 0 otherwise
Ear (H)	1 if diagnosis within ICPC-chapter "Ear" (H), 0 otherwise
Circulatory (K)	1 if diagnosis within ICPC-chapter "Circulatory" (K), 0 otherwise
Musculoskeletal (L)	1 if diagnosis within ICPC-chapter "Musculoskeletal" (L), 0 otherwise
Neurological (N)	1 if diagnosis within ICPC-chapter "Neurological" (N), 0 otherwise
Psychological (P)	1 if diagnosis within ICPC-chapter "Psychological" (P), 0 otherwise
Respiratory (R)	1 if diagnosis within ICPC-chapter "Respiratory" (R), 0 otherwise
Skin (S)	1 if diagnosis within ICPC-chapter "Skin" (S), 0 otherwise
Endocrine (T)	1 if diagnosis within ICPC-chapter "Endocrine, metabolic and nutritional" (T), 0 otherwise
Urology (U)	1 if diagnosis within ICPC-chapter "Urology" (U), 0 otherwise
Pregnancy (W)	1 if diagnosis within ICPC-chapter "Pregnancy, childbirth, family planning" (W), 0 otherwise
Female genital system (X)	1 if diagnosis within ICPC-chapter "Female genital system" (X), 0 otherwise
Male genital system (Y)	1 if diagnosis within ICPC-chapter "Male genital system" (Y), 0 otherwise
Industry:	
Agriculture	1 if patient is working in agriculture, forestry or fishing, 0 otherwise
Mining	1 if patient is working in mining or electricity, gas and water supply, 0 otherwise
Manufacturing	1 if patient is working in manufacturing, 0 otherwise
Construction	1 if patient is working in construction, 0 otherwise
Wholesale and retail	1 if patient is working in wholesale, retail trade, hotel or restaurant, 0 otherwise
Transport	1 if patient is working in transport, 0 otherwise
Financial	1 if patient is working in financial intermediation, real estate, renting and business activities, 0 otherwise
Public administration	1 if patient is working in public administration, 0 otherwise
Education	1 if patient is working in education, 0 otherwise
Health	1 if patient is working in health or social work, 0 otherwise
Physician characteristics	
Age_GP	Age of the GP
Male_GP	1 if the GP is male, 0 otherwise
Specialist_GP	1 if the GP with some specialization, 0 otherwise
List length	Number of patients on the list
Fixed wage	1 for GP with fixed salary, 0 otherwise
Municipality characteristics	
Large cities	1 if patient resident in large city, 0 otherwise
Other urban areas	1 if patient is resident in other urban area, 0 otherwise
Rural areas	1 if patient is resident in rural area, 0 otherwise
Index mortality	Index mortality (1 – 10, 10 for municipalities with the highest mortality)
Index unemployment	Index unemployment (1 – 10, 10 for municipalities with the highest unemployment)

Table 2: Descriptive statistics of the variables used in the regression analyses

Variable	All patients Mean (S.D)[a]	Males Mean (S.D)[a]	Females Mean (S.D)[a]
Patient characteristics			
Sick days	62.267 (77.443)	63.457 (79.235)	61.500 (76.257)
Male	0.392 (0.488)	-	-
Age	42.150 (11.781)	42.516 (11.944)	41.915 (11.668)
Years of education	12.354 (3.228)	12.027 (3.137)	12.566 (3.268)
Income/1000	2208.16 (2139.32)	2583.48 (2912.18)	1966.43 (1383.18)
Number of individuals in the household	2.724 (1.390)	2.715 (1.469)	2.730 (1.338)
Number of children < 6	0.248 (0.545)	0.234 (0.553)	0.256 (0.540)
Years of experience	19.539 (10.311)	21.949 (11.186)	17.986 (9.384)
Sick days in 2002	19.257 (40.356)	17.430 (38.581)	20.433 (41.417)
Working hours per week			
Working hours 4-19	0.127 (0.326)	0.053 (0.210)	0.174 (0.377)
Working hours 20-29	0.110 (0.312)	0.024 (0.152)	0.165 (0.371)
Working hours 30 +	0.763 (0.425)	0.923 (0.267)	0.661 (0.473)
Chapters in ICPC			
General and unspecified (A)	0.049 (0.211)	0.047 (0.201)	0.051 (0.323)
Blood (B)	0.004 (0.066)	0.004 (0.064)	0.004 (0.067)
Digestive (D)	0.046 (0.209)	0.056 (0.229)	0.040 (0.195)
Eye (F)	0.008 (0.087)	0.010 (0.098)	0.006 (0.079)
Ear (H)	0.007 (0.081)	0.007 (0.082)	0.006 (0.080)
Circulatory (K)	0.043 (0.202)	0.061 (0.239)	0.031 (0.173)
Musculoskeletal (L)	0.440 (0.496)	0.496 (0.499)	0.403 (0.491)
Neurological (N)	0.044 (0.206)	0.041 (0.197)	0.047 (0.211)
Psychological (P)	0.163 (0.370)	0.151 (0.358)	0.171 (0.377)
Respiratory (R)	0.062 (0.241)	0.062 (0.241)	0.062 (0.241)
Skin (S)	0.028 (0.166)	0.039 (0.193)	0.022 (0.146)
Endocrine (T)	0.013 80.115)	0.011 (0.104)	0.015 (0.122)
Urology (U)	0.007 (0.085)	0.008 (0.087)	0.007 (0.084)
Pregnancy (W)	0.061 (0.239)	0	0.100 (0.300)
Female genital system (X)	0.021 (0.144)	0	0.035 (0.183)
Male genital system (Y)	0.003 (0.057)	0.008 (0.090)	0
Industry			
Agriculture	0.008 (0.089)	0.013 (0.115)	0.005 (0.067)
Mining	0.016 (0.126)	0.031 (0.172)	0.007 (0.083)
Manufacturing	0.130 (0.239)	0.225. (0.415)	0.067 (0.259)
Construction	0.061 (0.238)	0.141 (0.348)	0.010 (0.095)
Wholesale and retail	0.154 (0.361)	0.152 (0.361)	0.154 (0.361)
Transport	0.082 (0.275)	0.136 (0.342)	0.048 (0.213)
Financial	0.092 (0.290)	0.094 (0.291)	0.092 (0.288)
Public administration	0.067 (0.249)	0.057 (0.232)	0.073 (0.260)
Education	0.087 (0.282)	0.049 (0.215)	0.112 (0.316)
Health	0.303 (0.459)	0.102 (0.303)	0.432 (0.495)
Physician characteristics			
Age_GP	47.520 (8.768)	47.733 (8.995)	47.384 (8.616)
Male_GP	0.721 (0.449)	0.815 (0.388)	0.660 (0.474)
Specialist_GP	0.613 (0.487)	0.602 (0.489)	0.620 (0.485)
List length	1334.67 (392.239)	1332.714 (395.362)	1335.939 (390.210)
Fixed wage	0.063 (0.243)	0.067 (0.249)	0.061 (0.239)
Municipality characteristics			
Large cities	0.583 (0.492)	0.568 (0.494)	0.594 (0.491)
Other urban areas	0.290 (0.454)	0.302 (0.459)	0.282 (0.450)
Rural areas	0.127 (0.332)	0.130 (0.337)	0.124 (0.330)
Index mortality	5.574 (2.131)	5.584 (2.149)	5.568 (2.120)
Index unemployment	6.117 (2.289)	6.122 (2.293)	6.113 (2.287)
Number of episodes	353,611	138,524	215,087
Number of physicians	3,709	3,691	3,703
Number of municipalities	415	415	414

[a] Means or proportions of variables, standard deviation of variables in parenthesis

Table 3: The average length of sickness absence per main chapter in ICPC

Diagnosis	All patients (n=353,611) Mean (S.D)[a]	Males (n=138,524) Mean (S.D)[a]	Females (N=215,087) Mean (S.D)[a]
General and unspecified (A)	48.177 (67.990)	48.476 (68.663)	47.998 (67.585)
Blood (B)	86.071 (92.285)	107.859 (98.050)	73.084 (86.092)
Digestive (D)	46.120 (66.727)	45.885 (66.436)	46.333 (66.994)
Eye (F)	49.307(65.843)	45.529 (68.256)	42.390 (67.526)
Ear (H)	68.022 (86.478)	73.922 (90.428)	63.987 (83.463)
Circulatory (K)	78.877 (88.272)	90.651 (90.722)	64.030 (82.733)
Musculoskeletal (L)	66.392 (79.418)	63.659 (77.601)	68.559 (80.764)
Neurological (N)	66.109 (82.785)	71.720 (86.431)	62.950 (80.493)
Psychological (P)	75.933 (84.049)	79.403 (85.371)	73.963 (83.226)
Respiratory (R)	28.766 (56.900)	32.597 (62.430)	26.288 (52.871)
Skin (S)	39.858 (64.262)	36.030 (59.819)	44.266 (68.768)
Endocrine (T)	74.122 (85.329)	81.598 (89.192)	70.634 (83.249)
Urology (U)	44.051 (68.485)	56.280 (77.228)	35.554 (60.267)
Pregnancy (W)	51.923 (45.510)	-	51.923 (45.510)
Female genital system (X)	54.725 (73.738)	-	54.725 (73.738)
Male genital system (Y)	56.478 (73.233)	56.478 (73.233)	-
All	62.267 (77.443)	63.458 (79.235)	61.500 (76.257)

[a] Means and standard deviation of variables in parenthesis

Table 4: Fixed part results of the random intercept model, where the length of sick leave (LSL) is the dependent variable

	Total sample	Male patients	Female patients
Patient characteristics			
Male	1.129** (0.328)	-	-
Age	1.102** (0.024)	1.517** (0.050)	1.023** (0.027)
Years of education	-0.819** (0.043)	-1.162** (0.070)	-0.659** (0.054)
Income/1000	-0.038** (0.006)	-0.043** (0.007)	-0.029* (0.012)
Number of individuals in the household	-1.395** (0.103)	-1.825** (0.159)	-0.902** (0.135)
Number of children < 6	3.321** (0.269)	0.564 (0.426)	5.158** (0.347)
Years of experience	-0.538** (0.026)	-0.783** (0.053)	-0.448** (0.032)
Sick days in 2002	0.109** (0.003)	0.120** (0.005)	0.100** (0.004)
Working hours per week: Base category: Working hours 4-19			
Working hours 20-29	-5.596** (0.523)	-11.061** (1.602)	-4.574** (0.551)
Working hours 30 +	-6.190** (0.404)	-11.922** (0.926)	-4.644** (0.453)
Chapters in ICPC: Base category: General and unspecified (A)			
Blood (B)	35.775** (2.000)	55.498** (3.330)	23.715** (2.490)
Digestive (D)	-3.358** (0.812)	-4.982** (1.288)	-2.212* (1.074)
Eye (F)	-7.208** (1.563)	-5.817** (2.301)	-8.507** (2.145)
Ear (H)	17.029** (1.662)	22.250** (2.668)	13.821** (2.120)
Circulatory (K)	23.757** (0.846)	32.968** (1.278)	11.235** (1.160)
Musculoskeletal (L)	15.585** (0.604)	12.479** (0.994)	18.384** (0.758)
Neurological (N)	16.458** (0.830)	21.062** (1.393)	13.902** (1.029)
Psychological (P)	27.116** (0.651)	29.447** (1.086)	25.880** (0.811)
Respiratory (R)	-21.662** (0.765)	-18.224** (1.257)	-23.763** (0.960)
Skin (S)	-9.150** (0.945)	-12.525** (1.411)	-4.948** (1.299)
Endocrine (T)	23.388** (1.231)	26.845** (2.184)	21.496** (1.485)
Urology (U)	-7.681** (1.582)	1.229 (2.532)	-14.656** (2.022)
Pregnancy (W)	12.856** (0.787)	-	11.141** (0.892)
Female genital system (X)	4.028** (1.048)	-	4.500** (1.119)
Male genital system (Y)	2.165 (2.304)	1.121 (2.459)	-
Industry: Base category: Manufacturing			
Agriculture	9.619** (1.464)	9.300** (1.841)	9.352** (2.470)
Mining	-9.645** (1.064)	-10.166** (1.270)	-8.092** (2.025)
Construction	8.006** (0.631)	9.925** (0.710)	5.163** (1.790)
Wholesale and retail	6.832** (0.496)	7.780** (0.701)	5.752** (0.749)
Transport	-1.644** (0.572)	-1.094 (0.721)	-3.257** (0.964)
Financial	2.210** (0.565)	3.616** (0.828)	0.519 (0.822)
Public administration	-6.988** (0.626)	-11.455** (0.991)	-5.359** (0.867)
Education	-2.786** (0.602)	-2.508* (1.096)	-3.782** (0.809)
Health	-5.750** (0.468)	-6.080** (0.807)	-6.256** (0.678)
Physician characteristics			
Age_GP	-0.002 (0.022)	0.033 (0.030)	-0.015 (0.027)
Male_GP	-0.578 (0.399)	-1.328* (0.620)	-0.063 (0.472)
Specialist	0.131 (0.392)	0.222 (0.541)	0.135 (0.482)
List length/100	-0.161** (0.051)	-0.151* (0.070)	-0.180** (0.063)
Fixed wage	-2.185** (0.823)	-1.893 (1.096)	-2.784** (1.012)
Municipality characteristics			
Municipality location: Base category: Large cities			
Other urban areas	-0.065 (1.132)	1.265 (1.193)	-0.975 (1.277)
Rural areas	-2.008 (1.112)	0.985 (1.239)	-3.501** (1.271)
Index mortality	0.250 (0.173)	0.432* (0.193)	0.122 (0.199)
Index unemployment	-0.162 (0.171)	0.096 (0.192)	-0.255 (0.197)
Constant	32.144** (2.075)	30.499** (2.894)	33.378** (2.513)
Number of episodes	353,611	138,524	215,087
Number physicians	3,709	3,691	3,703
Number municipalities	415	415	414

Note: Standard errors are in parenthesis. '*' and '**'represents significance level at the 5% and 1% level respectively.

Table 5: Random effects parameters for the random intercept models, where the length of sick leave (LSL) is the dependent variable

	All Patients	Males	Females
	Variance component (se)	Variance component (se)	Variance component (se)
Model I: Null model			
Municipalities (σ_m^2)	57.007 (5.898)	56.964 (7.050)	64.117 (7.398)
Physicians (σ_G^2)	36.584 (2.418)	33.826 (4.549)	46.686 (3.526)
Residuals (σ_i^2)	5913.608 (14.129)	6197.376 (23.798)	5717.174 (17.562)
Model II: includes patient-level covariates only			
Municipalities (σ_m^2)	50.017 (5.344)	46.122 (5.999)	60.103 (7.090)
Physicians (σ_G^2)	35.397 (2.301)	31.164 (4.222)	47.415 (3.415)
Residuals (σ_i^2)	5613.526 (13.413)	5804.369 (22.289)	5442.507 (16.718)
Model III: includes patient-level and GP-level covariates			
Municipalities (σ_m^2)	48.099 (5.216)	44.694 (5.905)	57.014 (6.865)
Physicians (σ_G^2)	35.284 (2.304)	31.216 (4.233)	47.390 (3.420)
Residuals (σ_i^2)	5613.513 (13.412)	5804.262 (22.289)	5442.543 (16.718)
Model IV: includes all three-level covariates			
Municipalities (σ_m^2)	48.205 (5.238)	43.074 (5.834)	56.873 (6.838)
Physicians (σ_G^2)	35.207 (2.301)	31.304 (4.236)	47.204 (3.411)
Residuals (σ_i^2)	5613.510 (13.412)	5804.323 (22.288)	5442.509 (16.718)
Number of episodes	353,611	138,524	215,087
Number physicians	3,709	3,691	3,703
Number municipalities	415	415	414

Note: Standard errors are in parenthesis.