Comparisons between five self-administered instruments predicting sick leaves in a 4-year follow-up

Per Lindberg[1, 2], PhD

Malin Josephson[3], PhD

Lars Alfredsson[4, 5], PhD

Eva Vingård[1, 3], MD, PhD, Prof

[1]Section for Personal Injury Prevention, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden.

[2] Centre for Musculoskeletal Research, University College of Gävle, Sweden

[3]Section of Occupational and Environmental Medicine, Department of Medical Sciences, Uppsala University, Sweden.

[4]Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.

[5]Centre of Public Health, Stockholm County Council, Stockholm, Sweden.

Correspondence to:
Per Lindberg,
University College of Gävle
Centre for Musculoskeletal Research
SE-801 76 Gävle
Sweden

E-mail: per.lindberg@hig.se; Phone: +46 26 648 759; +46 709 689 686

Abstract

Objectives This study aimed to explore and compare the ability of five instruments for self-rating to predict future sick leave rates.

Methods In three Swedish municipalities 2 252 employees completed a baseline questionnaire and were followed up for 4 years. Five health-oriented instruments for self-rating were used as potential predictors of the two outcome measures no sick leave at all, and one or more spells of long-term sick leave ≥ 28 days. Positive and negative predictive values as well as Cox proportional hazard ratios (denoted as RRs) adjusted for age and work type were calculated.

Results The instruments showed no statistical difference in predicting future sick leave for either of the sexes. For no sick leave RRs ranged between 1.27 and 1.52 (women), 1.35 and 1.61 (men); for long-term sick leave RRs ranged between 1.78 and 2.39 (women), 2.87 and 5.53 (men). However the best prediction of long-term sick leave for men, RR 5.53, 95% confidence interval (CI) 3.37–9.08, was significantly higher than the best prediction for women, RR 2.39, 95% CI 1.97–2.90.

Conclusion Prediction of long-term sick leave was better than that of no sick leave, and better among men than among women. There was a tendency for somewhat better prediction of future sick leave by multiple-question instruments, but single-question instruments can very well be used in predicting future sick leaves, and crude analyses stratified by sex can be used for screening purposes.

Keywords sickness absence; self-rated health; SRH; determinant; work ability

Introduction

Self-rated health (SRH) is among the most frequently assessed health perspectives in clinical and epidemiological studies. It is frequently used as an end-point but a large number of studies have demonstrated that a person's own appraisal of her/his health is a powerful predictor of future morbidity (Segovia et al. 1989; Bjorner et al. 1996; Kaplan et al. 1996; Moller et al. 1996; Huibers et al. 2004b) and mortality (Bjorner et al. 1996; Idler and Benyamini 1997; Benyamini and Idler 1999; Bardage et al. 2001; Burstrom and Fredlund 2001; Heidrich et al. 2002; Larsson et al. 2002; Af Sillén et al. 2005; Vuorisalmi et al. 2005). SRH is considered a cheap and convenient way of identifying risk groups and risk factors (Bjorner et al. 1996), useful not least in working life. It is noteworthy that, although studies of positive aspects of health have been recommended (Mackenbach et al. 1994; Manderbacka et al. 1998), to our knowledge only a few studies have analyzed SRH as a predictor of positive outcomes, such as remaining in work, recovery from fatigue, and survival among cancer patients (Lund and Borg 1999; Shadbolt et al. 2002; Huibers et al. 2004a).

In epidemiological studies SRH is often operationalized by a single question. There are a large variety of question-response combinations; however, most questions can be categorized into non-comparative, age-comparative, and time-comparative. The non-comparative question has been suggested as standard single-question measure (Bjorner et al. 1996; Eriksson et al. 2001). SRH may also be obtained by multiple questions, which have the advantage of increased reliability and validity and better ability to discriminate (Bjorner et al. 1996;

Sullivan et al. 2002; Bowling 2005). Internationally one frequently used scale is the 36-item Short Form (SF-36) Health Survey, in which 36 questions form eight sub-scales including a scale on general health (SF-36 GH) (Ware and Sherbourne 1992; Ware et al. 1993; Sullivan et al. 1995; Sullivan et al. 2002).

A concept closely related to health is functional ability, the ability to perform tasks and roles within a social context, e.g. to work. Self-assessment of work ability is one way of defining functional status from a working life perspective. Though, from a conceptual point of view functional ability should be distinguished from health (Bjorner et al. 1996), studies have shown self-assessed reduced work ability to be significantly associated with future sick leave (Reiso et al. 2001; Kujala et al. 2006), but not with “no sickness absence”(Kujala et al. 2006).

Short and valid instruments for scientific and screening purposes are sought after (Bowling 2005). From en occupational health standpoint it is of interest to identify target groups for intervention as well as groups with excellent health (no sick leave) to learn from. However the number of instruments for assessing and predicting health-related outcomes raises several questions: for instance, does the predictive precision differ between different instruments? Are there predictive differences between single- and multiple-question instruments? Do the predictive qualities differ between women and men? Are the instruments sensitive for both end-points of a health continuum, excellent and poor? Thus, the aim of this study was to explore and compare the ability of five health-oriented instruments for self-assessment to predict no sick leave as well as long-term sick leave.

Material

The present study forms part of the longitudinal study "Work and sustainable health in the public sector in Sweden", the HAKuL study (Vingård et al. 2005), which was launched in 1999–2000 in four county councils and local authorities of six municipalities in Sweden. In total 9 003 persons employed by these county councils and local authorities were asked to participate. The study started with a baseline postal questionnaire on individual factors, social situation, health, lifestyle, and work factors sent to all employees who were not on long-term sick leave for 3 months or longer at the time of the baseline questionnaire. The HAKuL study has been approved (No. 99-242) by the Ethics Committee at Karolinska Institutet, Stockholm, Sweden, in accordance with 1964 Declaration of Helsinki. All persons gave their informed consent prior to their inclusion in the study.

Study group

In three of the local authorities (in central and southern Sweden) taking part in the HAKuL study we had access to the employers' registers on all absences from work covering the period from the point the baseline questionnaire was filled in during year 2000 and the following 4 years. The baseline questionnaire was completed and returned by 2 306 individuals. After excluding 54 persons due to missing follow-up data, 2 252, 84% of the asked, remained for the analyses. The sample consisted of 79% women with median age 45.5 years (men: 49 years). The proportion with university training was the same for women and men, around 30%, but men in general had a lower educational level. This was mirrored in the work characteristics, with a larger proportion of men than of women having

mainly physical work tasks and a larger proportion of women compared with men having mixed work tasks (Table 1).

Table 1 about here

Methods

Outcome measures

In this study the outcome measures "no sick leave" was defined as not having had one, 1, single day of sick leave and "long-term sick leave" as having had one or more spells of sick leave ≥ 28 days during 4 years of follow-up. The cut-off point at 28 days was chosen as, according to Swedish legislation at the time for the study, the employer had to produce a plan for rehabilitation when an employee had been on sick leave for 28 days, thus denoting this to be a spell of sick leave of considerable length, eventually in need of special attention to facilitate return to work. Because of access to the employers' registers it was possible to accurately calculate length of sick leave spells and days at risk for the risk analyses.

Potential predictors of sick leave

From the baseline questionnaire five health-oriented instruments for self-assessment, described below, were chosen as possible predictors and their answer categories were dichotomized in two ways as two outcomes were analyzed. The choice of cut-off points for the "excellent" and "poor" predictor levels was made as recommended in the literature or as conventionally made in population studies, using the upper and lower quartile for indices and a more semantic, "what is reasonable" approach for single items.

1. Global SRH

We used the non-comparative global health question: “In general, how would you describe your health? As excellent/very good/good/fair/poor?” The response categories “excellent” and “very good” were called “excellent Global SRH” and compared with the less good categories (good/fair/poor) to predict no sick leave. The response categories “fair” and “poor” were termed “poor Global SRH” and compared with the better than poor categories (excellent/very good/good) to predict long-term sick leave.

2. SF-36 GH

In addition to global SRH (as above), the SF-36 GH scale includes the following four statements: (1) I seem to get sick a little easier than other people; (2) I am as healthy as anybody I know; (3) I expect my health to get worse; and (4) My health is excellent. Five response choices are available, “definitely true”, “mostly true”, “don’t know”, “mostly false” and “definitely false”. The values of the five items are summed and transformed to a 0–100 scale (Sullivan et al. 2002). Our cut-offs were set according to the quartiles in the HAKuL baseline (n=7 533). “Excellent SF-36 GH” (= highest quartile, score ≥ 87.5) versus less good SF-36 GH (< 87.5) was used to predict no sick leave and “poor SF-36 GH” (= lowest quartile, score ≤ 57) versus the better than poor (> 57) to predict long-term sick leave.

3. Work ability index

The work ability index (WAI) (Ilmarinen et al. 1991; Tuomi et al. 1991; Tuomi et al. 1997) is determined on the basis of the answers to ten questions covering seven

items, viz. (1) current work ability compared with the lifetime best; (2) work ability in relation to job demands; (3) number of diagnosed current diseases; (4) estimated work impairment due to diseases; (5) number of days on sick leave last 12 months; (6) own prognosis of work ability 2 years from now; and (7) mental resources in general. A sum score is calculated and the respondents are classified into poor, moderate, good, and excellent (Tuomi et al. 1998). In this study we used a slightly modified version of WAI (Lindberg 2006). The correlation between the original and the modified WAI categories was good (Cohen's kappa 0.84). "Excellent WAI"(WAI category: excellent) versus less good WAI (good/moderate/poor) was used to predict no sick leave and "poor WAI" (WAI categories: poor and moderate) versus better than poor WAI (good/excellent) to predict long-term sick leave.

4. Work-ability-in-2-years

Item six of the WAI index reads, "Do you believe that, from your health perspective, you will be able to do your current job 2 years from now?", with answer categories "relatively certain", "not certain" and "unlikely"(Tuomi et al. 1998). Considering the predictive capabilities of global SRH we hypothesized that the prognostic question on work ability could have similar qualities. "Excellent Work-ability-in-2-years" (response category: relatively certain) versus poor Work- ability-in-2-years (not certain/unlikely) was used to predict no sick leave and "poor Work-ability-in-2 years" (not certain and unlikely) versus excellent Work-ability-in-2-years (relatively certain) to predict long-term sick leave.

5. Health-for-working

Within the HAKuL study (Vingård et al. 2005) a two-item scale titled, "Health-for-working", was constructed combining "Global SRH" and "Work-ability-in-2-years". On the basis of considering good health as a sense of wellbeing and as signifying the ability to act (Nordenfelt 1987), Health-for-working intends to capture the relationship between health and work. The scores for Global SRH (excellent = 1 point – poor = 5 points) were summed with weighted scores of Work-ability-in-2-years (relatively certain = 1.5 points; not certain = 2 points; unlikely = 5 points). The sum score was divided into five Health-for-working categories (excellent = 2.5–3 points; good = 3.5 points; moderate = 4–4.5 points; failing = 5–5.5 points; poor = 6–10 points). This instrument has so far not been validated. "Excellent Health-for-working" (Health-for-working category: excellent) versus the less good categories (good/moderate/failing/poor) was used to predict no sick leave and "poor Health-for-working" (category: poor) versus the better than poor categories (failing/moderate/good/excellent) to predict long-term sick leave.

Statistical analyses

Predictive abilities of the five instruments were separately analyzed for the two outcomes and stratified by sex. Positive and negative predictive values were calculated, where the positive predictive value is the proportion of exposed individuals that became cases and negative predictive value is the proportion of non-exposed individuals that did not became a cases (Altman and Bland 1994). Cox proportional hazard regression model was used giving hazard ratios (denoted as rate ratios (RRs)) together with 95% confidence intervals (CIs) for the

associations between individual baseline measures and the outcomes. The calculated hazard ratios were adjusted for sex, age and work type. The work types were classified according to the WAI instrument (Tuomi et al. 1998) into mainly physical (e.g. home-based personal care, and craftsmen), mainly mental (e.g. teachers and managers) or mixed (e.g. nurses and assistant nurses). During the 4 years 25% of the employees ended their employment, which was considered when calculating days at risk. Cox regressions are dependent on an "event". To comply with this, the calculations of RRs for no sick leave, which is a "non-event", had to be converted to RRs for having ≥ 1 day of sick leave, giving hazards < 0 for predictors rated as excellent. The calculated RRs were then inverted in order to give the RRs for (the chance of) having no sick leave. This procedure also intended to facilitate the comparison between excellent and poor predictors. For the statistical analyses SPSS 13.0 was used.

Non-respondents and internal missing values

There were in general few internal missing values (0.4–2.2%). However, for the WAI instrument, based on the largest number of variables, 172 respondents (8%) were not able to be WAI-classified because of one or more missing values. Analyses revealed that compared with the WAI-classified respondents a somewhat greater proportion of those with missing values reported poor global health at baseline and developed long-term sick leave during the follow-up (not shown).

Results

With the exception of “Excellent Work-ability-in-2-years”, in all instruments the same proportion of women as of men rated their health/work ability as good while a larger proportion of women rated their health/work ability as poor at baseline. At follow-up 4 years later 18% of women and 32% of men had not been on sick leave one single day while 28% of women and 18% of men had been on long-term sick leave at least once (Table 2).

Table 2 about here

Prediction

The positive predictive values of the five instruments showed a wider range for the outcome no sick leave among women, 17-29%, than among men, 34-41%. For the outcome long-term sick leave this was reversed, women: 38-43% and men: 38-55%. The negative predictive values showed overall higher values and the ranges were somewhat more narrow, for no sick leave, women: 80-86% and men: 69-76%, and for long-term sick leave, women: 75-79 and men:75-88%.

The RRs of the five instruments did not differ significantly in predicting no sick leave (Table 3) and were almost identical for women and men.

Table 3 about here

For predicting long-term sick leave, there were no statistically significant differences between the instruments within respective sex (Table 4). For women,

the predictions of the five instruments were almost identical, poor WAI showing the strongest association (RR 2.39, 95% CI 1.97–2.90). For men, the RRs for the five instruments were on a distinctly higher level and showed a wider range, being strongest for poor Health-for-working (RR 5.53, 95% CI 3.37–9.08). Two of the instruments, Work-ability-in-2-years and Health-for-working, had significantly higher associations for men rating poor than for women rating poor.

Table 4 about here

Discussion

Five instruments for self-assessment were analyzed with regard to their ability to predict future sick leave. The instruments, showing similar power to predict, predicted long-term sick leave better than no sick leave and predicted long-term sick leave somewhat better for men than for women. Similar findings have recently been reported by Kujala et al, who in a one-year follow-up investigating a slightly modified WAI found WAI to be a practical tool for predicting long-term sick leave (>9 days), but limited in predicting no sick leave among young employees (Kujala et al. 2006).

One strength of this study is the reliable data on sick leave as we had access to the employers' registers of all absences from work and could at an individual level calculate accurate times at risk and outcome categories. Another strength is that there were few internal missing values except with the WAI instrument. As the respondents whose data were missing in the WAI analyses had poorer health at baseline and more long-term sick leave during the follow-up, it is likely that the

RRs for WAI in this study are slightly underestimated, but not so much that they would greatly change the results.

Since the study population was not a random sample of the general population there may be some limitations in how to generalize our findings. However, as our intention was to explore if there were any differences between five self-administered instruments in their ability to predict sick leave, it is likely that the results, that the instruments were equally good, would be similar in a general population. In contrast, if the aim had been to compare the absolute levels of the associations it would not have been possible to generalize our findings as the prevalence of health/workability and sick leave vary between different groups in society.

As we found a marked lower discriminating ability for excellent compared with poor baseline assessments, we tested with cut-offs at the upper and lower 15th percentile to explore whether this would show stronger associations. These calculations were possible for SF-36 GH, WAI and Health-for-working. The RRs turned out almost identical to the ones presented in Tables 3 and 4, except for excellent WAI, which showed a tendency for a stronger association with no sick leave (RR 1.82, 95% CI 1.54–2.13 for women and men together) (not shown). This was an expected outcome as the predictor excellent WAI now included the best 15% instead of the best 40%. In order to achieve better discriminating ability for response alternatives at the positive end of the scale, it is obvious that the scales have to be further developed, e.g. to include more categories or other wording. However this is not an easy task. The difficulties to assess positive

health have been addressed by Kemm (Kemm 1993) who mentioned the paucity of language to be one ground for these difficulties.

As three of the predictors included the prognostic question on work ability in 2 years we hypothesized that the predictive precision would be even better at a 2-year follow-up. Performing the same analyses as described here did not confirm this hypothesis; the results (not shown) were almost identical, which could indicate that the yearly incidence of sick leave was fairly constant.

A greater proportion of women than of men scored low (= poor) on the investigated instruments and experienced long-term sick leave during the follow-up. However, the relative risk for developing long-term sick leave was markedly higher for men—statistically significant for two instruments. In other words, at the "poor" end-point the instruments seemed to predict future long-term sick leave better among men than among women and thus give stronger point estimates. It has been shown that women's SRH assessments are based on a wider range of health-related and non-health-related factors than are men's (Benyamini et al. 2000). This difference in self assessment of health may cause a slight dilution of the association among women and explain the gender differences in the prediction of future sick leave patterns found in this study.

The described associations decreased only to a minor extent from crude to adjusted models. To check that we had not lost any "hidden" information we also performed analyses stratified in four age groups (20-35, 36-45, 46-55, and 56-65 years) and by the three work types (not shown). With one exception these also did

not show any statistically significant differences, neither between the strata for a specific instrument nor between the instruments within each strata, nor between crude and adjusted analyses. This was true both for women and for men. The only exception was for the predictor Poor Health-for-working, when adjusted for work type, there was a significant difference for men belonging to the age groups 36-45 years (RR 54.99, 95% CI 9.46-319.57) and 46-55 years (RR 3.94, 95% CI 1.67-9.30). This difference can probably be explained by the few exposed cases in the 36-45 group. Taken together the findings in this study point out that the instruments are all rather stable within respective sex despite age and work type. Thus we can conclude that also the crude measurements, when stratified by sex, are fairly reliable to use when more sophisticated analyses are not possible, e.g. for screening purposes in occupational health care.

One of the research questions concerned the use of single questions versus multiple questions. A similar approach had a study about prediction of mortality and healthcare utilization where no difference between single- and multi-item measures was found (DeSalvo et al. 2005). We found tendencies for differences between the point estimates, but the confidence intervals overlapped. It has been suggested that single-question items are better suited as end-points while multiple-question items are better for independent variables (Bowling 2005). We will not argue against this standpoint but according to our results, if space is limited a single question on global health as determinant is still a reasonable choice.

Conclusion

We could not find any statistically significant differences between the abilities of the five instruments to predict future sick leave rates. Even if there was a tendency for stronger associations with multiple-question instruments, the single questions "Global SRH" and "Work-ability-in-2-years" can do very well, and crude analyses stratified by sex can be used for screening purposes. These results can be of importance when wanting to limit the length of a questionnaire. The predictive ability of the instruments worked in both directions, but predicted long-term sick leave better than no sick leave and better among men than among women. This points out instrumental limitations and calls for further development of the instruments to render them more sensitive for measures of the positive end-point in a health continuum.

Acknowledgements

The authors would like to thank statistician Stefan Stark of Uppsala University for preparing the data sets for the analyses. The HAKuL study was funded by a grant from AFA–Labour Market Insurance. There are no conflicts of interest.

REFERENCES

Af Sillén U, Nilsson J-Å, Månsson N-O, Nilsson P (2005) Self-rated health in relation to age and gender: influence on mortality risk in the Malmö Preventive Project. Scand J Public Health 33:183-189

Altman DG, Bland JM (1994) Statistics Notes: Diagnostic tests 2: predictive values. BMJ 309:102

Bardage C, Isacsson D, Pedersen N (2001) Self-rated health as a predictor of mortality among persons with cardiovascular disease in Sweden. Scand J Public Health 29:13-22

Benyamini Y, Idler E (1999) Community studies reporting association between self-rated health and mortality. Research on Aging 21:392-301

Benyamini Y, Leventhal EA, Leventhal H (2000) Gender Differences in Processing Information for Making Self-Assessments of Health. Psychosom Med 62:354-364

Bjorner J, Kristensen T, Orth-Gomér K, Tibblin G, Sullivan M, Westerholm P (1996) Self-rated health - a useful concept in research, prevention and clinical medicine. In. Forskningsrådsnämnden (Swedish Council for Planning and Coordination of Research), Stockholm, Sweden, p 144

Bowling A (2005) Just one question: If one question works, why ask several? J Epidemiol Community Health 59:342-345

Burstrom B, Fredlund P (2001) Self rated health: Is it as good a predictor of subsequent mortality among adults in lower as well as in higher social classes? J Epidemiol Community Health 55:836-840

DeSalvo KB, Fan VS, McDonell MB, Fihn SD (2005) Predicting Mortality and Healthcare Utilization with a Single Question. Health Services Research 40:1234-1246

Eriksson I, Undén A-L, Elofsson S (2001) Self-rated health. Comparisons between three different measures. Results from a population study. International Journal of Epidemiology 30:326-333

Heidrich J, Liese AD, Lowel H, Keil U (2002) Self-Rated Health and its Relation to All-Cause and Cardiovascular Mortality in Southern Germany. Results from the MONICA Augsburg Cohort Study 1984-1995. Annals of Epidemiology 12:338-345

Huibers MJH et al. (2004a) Predictors of outcome in fatigued employees on sick leave: Results from a randomised trial. Journal of Psychosomatic Research 57:443-449

Huibers MJH, Kant IJ, Knottnerus JA, Bleijenberg G, Swaen GMH, Kasl SV (2004b) Development of the chronic fatigue syndrome in severely fatigued employees: predictors of outcome in the Maastricht cohort study. J Epidemiol Community Health 58:877-882

Idler E, Benyamini Y (1997) Self-rated health and mortality: A review of twenty-seven community studies. Journal of Health and Social Behavior 38:21-37

Ilmarinen J, Tuomi K, Eskelinen L, Nygård C, Huuhtanen P, Klockars M (1991) Summary and recommendations of a project involving cross-sectional and follow-up studies on the aging worker in Finnish municipal occupations (1981-1985). Scand J Work Environ Health 17:135-141

Kaplan G et al. (1996) Perceived Health Status and Mobidity and Mortality: Evidence from Kuopio Ischaemic Heart Disease Risk Factor Study. Int J Epidemiol 25:259-265

Kemm JR (1993) Towards an epidemiology of positive health. Health Promotion International 8:129-1334
Kujala V, Tammelin T, Remes J, Elina V, Ek E, Laitinen J (2006) Work ability index of young employees and their sickness absence during the following year. Scand J Work Environ Health 32:75-84
Larsson D, Hemmingsson T, Allebeck P, Lundberg I (2002) Self-rated health and mortality among young men: what is the relation and how may it be explained. Scand J Public Health 30:259-266
Lindberg P (2006) The work ability continuum. Epidemiological studies of factors promoting sustainable work ability. In: Department of Clinical Neuroscience, Section of Personal Injury Prevention. Karolinska Institutet, Stockholm
Lund T, Borg V (1999) Work environment and self-rated health as predictors of remaining in work 5 years later among danish employees 35-59 ears of age. Experimental Aging Research 25:429-434
Mackenbach JP, van den Bos J, Joung IMA, van de Mheen H, Stronks K (1994) The determinants of excellent health: different from the determinants of ill-health? Int J Epidemiol 23:1273-1281
Manderbacka K, Lahelma E, Martikainen P (1998) Examining the continuity of self-rated health. Int. J. Epidemiol. 27:208-213
Moller L, Kristensen TS, Hollnagel H (1996) Self rated health as a predictor of coronary heart disease in Copenhagen, Denmark. J Epidemiol Community Health 50:423-428
Nordenfelt L (1987) On the nature of health. An action-theoretic approach. Kluwer Academic Publishers, Dordrecht
Reiso H, Nygård JF, Brage S, Gulbrandsen P, Tellnes G (2001) Work ability and duration of certified sickness absence. Scand J Public Health 29:218-225
Segovia J, Bartlett RF, Edwards AC (1989) An empirical analysis of the dimensions of health status measures. Social Science & Medicine 29:761-768
Shadbolt B, Barresi J, Craft P (2002) Self-rated health as a predictor of survival among patients with advanced cancer. J Clin Oncol 20:2514-2519
Sullivan M, Karlsson J, Taft C (2002) SF-36 Hälsoenkät: Svensk Manual och Tolkningsguide (SF-36 Health Survey: Swedish Manual and Interpretation), 2nd edn. Sahlgrenska University Hospital, Gothenburg
Sullivan M, Karlsson J, Ware J, John E. (1995) The Swedish SF-36 Health Survey--I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. Social Science & Medicine 41:1349-1358
Tuomi K, Ilmarinen J, Eskelinen L, Järvinen E, Toikkanen J, Klockars M (1991) Prevalence and incidence rates of diseases and work ability in different work categories of municipal occupations. Scand J Work Environ Health 17:67-74
Tuomi K, Ilmarinen J, Jahkola A, Katajarinne L, Tulkki A (1998) Work Ability Index, 2nd edn. Finnish Institute of Occupational Health, Helsinki
Tuomi K et al. (1997) Summary of Finnish research project (1981-1992) to promote the health and work ability of aging workers. Scand J Environ Health 23:66-71

Ware JE, Sherbourne CD (1992) The MOS 36-item short form health survey (SF-36). I. Conceptual framework and item selection. Medical Care 30:473-481

Ware JE, Snow K, Kosinski M, Gandek B (1993) SF-36 Health Survey Manual and Interpretation Guide. New England Medical Center, The Health Institute, Boston, MA

Vingård E et al. (2005) Long-term sick-listing among women in the public sector and its associations with age, family, lifestyle and work. A three year follow up study. Scandinavian Journal of Public Health 33:370-375

Vuorisalmi M, Lintonen T, Jylha M (2005) Global self-rated health data from a longitudinal study predicted mortality better than comparative self-rated health in old age. Journal of Clinical Epidemiology 58:680-687

Table 1 Descriptive data of the study group (N=2 252).

Characteristics	Women (n=1 786)	Men (n=466)
Age (years)		
median value	45.5	49
range	20–65	21–65
Education		
compulsory school	24%	36%
secondary school	32%	25%
university	31%	32%
other training	14%	7%
*Work characteristics**		
mainly physical	26%	46%
mixed physical and mental	43%	20%
mainly mental	31%	34%

*Classification according to the Work Ability Index (Tuomi et al. 1998).

Table 2 Self reported health assessments at baseline and proportion of subjects that had no sick leave or at least one spell of sick leave lasting more than 28 days, respectively, over the 4-year observation period.

Variables	Women (n=1 786)	Men (n=466)
Potential predictors		
Excellent Global health	37%	38%
Poor Global health	25%*	18%*
Excellent SF-36 GH[†]	21%	22%
Poor SF-36 GH	27%*	22%*
Excellent WAI[‡]	39%	43%
Poor WAI	25%*	18%*
Excellent Work-ability-in-2-years	77%*	85%*
Poor Work-ability-in-2-years	23%*	15%*
Excellent Health-for-working	11%	13%
Poor Health-for-working	14%*	10%*
Cumulative sick leave at 4-year follow-up		
Having had no sick leave at all	18%*	32%*
Having had ≥ 1 sick leave spell ≥ 28 days	28%*	18%*

*Statistically significant at $p<0.05$ for the difference between the proportion of women and men.
[†]SF-36 GH = 36-item Short Form (SF-36) Health Survey, general health scale
[‡]WAI = work ability index.

Table 3 No sick leave. The abilities of five instruments for self-assessment to predict no days of sick leave during the 4-year follow-up expressed as positive and negative predictive values and as crude and adjusted hazard ratios (RRs) with 95% confidence intervals (CIs).

Predictors	Number of exposed cases	PV^+ %	PV^- %	RR (95% CI), crude	RR (95% CI), adjusted for age and work type
Women (n=1 786)					
Excellent Global SRH	152	23	86	1.39 (1.25–1.54)	1.41 (1.27–1.59)
Excellent SF-36 GH	95	26	85	1.47 (1.30–1.69)	1.49 (1.32–1.72)
Excellent WAI	141	22	85	1.49 (1.33–1.67)	1.52 (1.37–1.69)
Excellent Work-ability-in-2-years	227	17	80	1.25 (1.10–1.41)	1.27 (1.12–1.45)
Excellent Health-for-working	53	29	84	1.52 (1.27–1.82)	1.52 (1.27–1.82)
Men (n=466)					
Excellent Global SRH	73	41	74	1.39 (1.10–1.75)	1.54 (1.20–1.96)
Excellent SF-36 GH	41	40	70	1.28 (0.96–1.69)	1.35 (1.02–1.82)
Excellent WAI	77	40	74	1.45 (1.15–1.82)	1.54 (1.22–1.92)
Excellent Work-ability-in-2-years	131	34	76	1.54 (1.15–2.04)	1.61 (1.20–2.17)
Excellent Health-for-working	24	41	69	1.16 (0.82–1.67)	1.23 (0.87–1.75)

* SRH = self-rated health
† SF-36 GH = 36-item Short Form (SF-36) Health Survey, general health scale
‡ WAI = work ability index.

Table 4 Long-term sick leave. The abilities of five instruments for self-assessment to predict one or more sick leave spells ≥ 28 days) during the 4-year follow-up expressed as positive and negative predictive values and as crude and adjusted hazard ratios (RRs) with 95% confidence intervals (CIs).

Predictors	Number of exposed cases	PV^+ %	PV^- %	RR (95% CI), crude	RR (95% CI), adjusted for age and work type
Women (n=1 786)					
Poor Global SRH	190	43	78	2.26 (1.89–2.72)	2.02 (1.68–2.43)
Poor SF-36 GH	196	42	78	2.23 (1.86–2.67)	2.06 (1.72–2.48)
Poor WAI	176	44	79	2.66 (2.20–3.22)	2.39 (1.97–2.90)
Poor Work-ability-in-2-years	151	38	75	2.06 (1.70–2.50)	1.78 (1.46–2.17)
Poor Health-for-working	107	43	75	2.34 (1.88–2.90)	1.98 (1.59–2.46)
Men (n=466)					
Poor global SRH	32	38	87	3.19 (2.04–4.97)	2.87 (1.84–4.50)
Poor SF-36 GH	38	38	88	3.87 (2.49–6.01)	3.83 (2.46–5.95)
Poor WAI	31	39	88	4.31 (2.72–6.83)	3.80 (2.36–6.11)
Poor work ability in 2 years	30	38	75	4.88 (3.09–7.70)	4.53 (2.83–7.26)
Poor Health-for-working	24	55	87	6.29 (3.88–10.20)	5.53 (3.37–9.08)

* SRH = self-rated health
† SF-36 GH = 36-item Short Form (SF-36) Health Survey, general health scale
‡ WAI = work ability index