BMJ Open

Self-reported health problems and obesity predict sickness absence during a 12-month follow-up: a prospective cohort study in 21 608 employees from different industries

Minna Pihlajamäki,[1,2] Jukka Uitti,[1,3] Heikki Arola,[2] Jyrki Ollikainen,[4] Mikko Korhonen,[4] Tapio Nummi,[4] Simo Taimela [ID] [5,6]

ABSTRACT

Objectives To study whether self-reported health problems predict sickness absence (SA) from work in employees from different industries.

Methods The results of a health risk appraisal (HRA) were combined with archival data of SA of 21 608 employees (59% female, 56% clerical). Exposure variables were self-reported health problems, labelled as 'work disability (WD) risk factors' in the HRA, presence of problems with occupational well-being and obesity. Age, socioeconomic grading and the number of SA days 12 months before the survey were treated as confounders. The outcome measure was accumulated SA days during 12-month follow-up. Data were analysed separately for males and females. A Hurdle model with negative binomial response was used to analyse zero-inflated count data of SA.

Results The HRA results predicted the number of accumulated SA days during the 12-month follow-up, regardless of occupational group and gender. The ratio of means of SA days varied between 2.7 and 4.0 among those with 'WD risk factors' and the reference category with no findings, depending on gender and occupational group. The lower limit of the 95% CI was at the lowest 2.0. In the Hurdle model, 'WD risk factors', SA days prior to the HRA and obesity were additive predictors for SA and/or the accumulated SA days in all occupational groups.

Conclusion Self-reported health problems and obesity predict a higher total count of SA days in an additive fashion. These findings have implications for both management and the healthcare system in the prevention of WD.

Strengths and limitations of this study

► Our study is based on prospectively collected extensive data from various fields of industries and occupations.
► The coverage, accuracy and consistency of the registry-based sickness absence outcomes is superior to self-reports.
► Our advanced statistical model is able to control key potential confounders.
► Generalisations can only be made to a working population.

screening questionnaires have shown predictive value for identifying individuals with an increased risk of sickness absence (SA) or work disability (WD) due to health issues.[2–5]

In the present study, we used a health risk appraisal (HRA), which is widely used in Finland and The Netherlands as a part of preventive occupational health services (OHS) to recognise employees at WD risk and to target interventions for those in need. The HRA was able to identify employees with a high number of SA days in an earlier study.[6] The previous study population (n=1341) were mainly blue collars (61%) and males (88%) from the construction industry. Age, gender, occupational grade and the self-assessment of future work ability were strong determinants of SA.[6]

SA is a complex and multifactorial phenomenon determined by personal, sociodemographical, lisfestyle-related and health-related factors as well as organisational determinants, healthcare management and legislation.[7 8] Self-reported health problems predict SA and prolonged return to work.[6 9] The key psychosocial predictors of SA include individuals' own perceptions of health and work ability.[10–13] On the other hand, SA serves as a measure of

INTRODUCTION

The main goals of health surveillance are to prevent work-related illnesses, to support workers' health and work ability and to reduce absenteeism as defined in International directives (Council Directive 83/391/EEC) and International Labour Office's guidelines (ILO).[1] Screening questionnaires are used as a part of targeted health surveillance to identify workers at risk. Some

Table 1 Criteria for classifying employees into the HRA categories

Topic	Criteria
Work disability risk: at least one of the topics below	
Doubt of work ability	Self-rated future work ability: uncertain of own ability ('uncertain'), or quite sure ('not able') not being able to continue in the current job due to health reasons.
Impairment due to musculoskeletal problems at work	Numerical rating scale (0–10) score ≥5.
Pain hampering work	At least moderate pain that affects working ability at minimum three times a week.
Sleep problems	Problems in falling asleep or night awakenings AND daytime sleepiness daily or almost daily.
Depressive symptoms	DEPS score ≥11.
Work-related constant fatigue	Feeling being squeezed empty.
Work-related constant stress	Feeling tense, strained, nervous and/or anxious because work-related issues are on one's mind all the time.
Health risks: at least one of the topics below	
Weight problems*	BMI ≥30 or ≤18.5.
Diabetes risk	Diabetes risk score ≥11.
Excess use of alcohol	Males ≥350 mL/week, females ≥240 mL/week (expressed as absolute alcohol).
Some symptoms: at least one of the topics below	
Impairment due to musculoskeletal problems at work	Numerical rating scale (0–10) score=4.
Some depressive symptoms	DEPS score between 8 and 10.
Some sleep problems	Problems in falling asleep or night awakenings AND daytime sleepiness 3–5 times a week.
A chronic disease	Self-reported chronic diseases.
Symptoms	Self-reported symptoms.
Lifestyle issues: at least one of the topics below	
Smoking	Smoking=yes.
Physical inactivity	No physical activity during leisure time nor while commuting to work.
Overweight*	BMI between 25 and 30.
No findings	
Previous criteria are not met	

*Overweight and weight problems were not included in the HRA category but analysed separately in the fully adjusted model.
BMI, body mass index; DEPS, Depression Scale; HRA, health risk appraisal.

Table 2 Criteria for problems with occupational well-being

Feature	Criteria
Insufficient job control	High 'job demands' AND (low 'decision authority' OR low result concerning 'job contents').
Work-life conflict	The low result in 'work-life balance'.
Strain due to rewarding	Any of the individual questions concerning rewarding (meaningfulness of work; appreciation; income; career opportunities) in a category 'causes very much strain'.
Lack of social support	Bullying at workplace OR (no support from line manager AND no teamwork).
Overloaded	Always feeling squeezed empty due to work OR always feels stress.
Dissatisfied	Seldom enjoys life OR never content with the present job.

to health and lifestyle risk factors, the HRA used in our study includes questions modified from the job demand-control-social support (JDCS) model, the effort-reward imbalance (ERI) model, the work-life conflict (WLC) theories, level of stress and work satisfaction.[17–19]

Recent studies have suggested that obesity may be a risk factor for SA,[20–22] and that prevention of obesity may be cost-effective.[23] In some studies, body mass index (BMI) has been a predictor of SA in females but not in males.[20 22] It has been estimated that obesity is associated with an
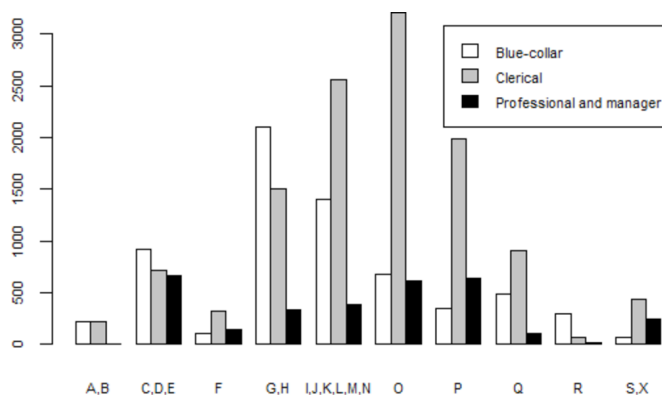
Figure 1 The distribution of the responses by occupational group and standard industrial classification by statistics Finland. A=agriculture, forestry and fishing; B=mining and quarrying; C=manufacturing; D=electricity, gas, steam and air conditioning supply; E=water supply; sewerage, waste management and remediation activities; F=construction; G=wholesale and retail trade; repair of motor vehicles; H=transportation and storage; I=accommodation and food service activities; J=information and communication; K=financial and insurance activities; L=real estate activities; M=professional, scientific and technical activities; N=administrative and support service activities; O=public administration and defence; compulsory social security; P=education; Q=health and social work activities; R=arts, entertainment and recreation; S=other service activities; X=industry unknown.

health in the working population when health is understood as a mixture of social, psychological and physiological functioning.[14 15] Multiple studies have shown that psychosocial work factors and work characteristics predict ill health and eventually WD.[16] Therefore, besides questions related

Table 3 The prevalence of work disability (WD) risk factors, problems with occupational well-being (OWB) and obesity and characteristics of the distribution of the number of sickness absence (SA) days by gender, occupational group (OG) and age

OG	Gender	Age	Subjects, N	WD risks (%) One	Two or more	Problems with OWB (%)	Obesity (BMI >30) (%)	Days on sickness absence % with zero SA	Median	Upper quartile	Mean of all values	Mean of non-zero values
BC	Female	≤30	701	19	13	25	13	29	5	13.0	12.7	17.9
		30–40	829	20	14	25	17	24	5	15.0	13.8	18.1
		40–50	1020	22	10	20	21	31	4	13.0	13.7	19.9
		50–60	1067	21	17	22	19	28	5	16.0	15.1	21.1
		>60	289	34	22	21	16	40	2	12.0	12.1	20.5
		All	3906	22	14	23	18	29	4	15	13.8	19.5
	Male	≤30	475	12	4	16	12	34	3	9	8.5	12.8
		30–40	702	16	6	19	20	29	3.5	10	9.7	13.6
		40–50	684	16	7	16	20	35	3	11	11.6	17.9
		50–60	659	20	12	17	24	38	2	11	11.5	18.7
		>60	177	25	13	11	24	53	0	8	9.9	21.1
		All	2697	17	8	17	20	35	3	10	10.4	16.1
	All		6603	28	12	20	19	32	4	12	12.4	18.2
C	Female	≤30	577	15	6	16	5	45	2	6	6.5	11.7
		30–40	1543	19	8	17	13	40	2	8	7.5	12.6
		40–50	2036	19	9	19	18	44	2	7	7.9	14.0
		50–60	2482	19	13	19	19	48	1	7	8.0	15.5
		>60	696	23	17	15	20	54	0	5	7.7	16.9
		All	7334	19	11	18	16	46	1	7	7.7	14.2
	Male	≤30	391	8	2	6	8	67	0	2.5	3.8	11.5
		30–40	1143	11	4	14	14	60	0	3	3.7	9.2
		40–50	1179	13	6	13	18	65	0	3	4.4	12.3
		50–60	1347	14	9	12	19	67	0	3	5.3	15.9
		>60	495	17	8	10	16	72	0	2	5.3	19.2
		All	4555	13	6	12	16	65	0	3	4.5	13.0
	All		11889	17	9	16	16	53	0	5	6.5	13.8
P/M	Female	≤30	60	20	0	10	10	58	0	3	3.8	9.1
		30–40	349	13	7	13	9	54	0	4	5.0	10.8
		40–50	485	18	6	13	12	56	0	4	4.5	10.1
		50–60	414	19	8	14	18	53	0	5	6.0	12.8
		>60	139	23	12	15	17	61	0	4	6.5	16.8
		All	1447	18	7	13	13	55	0	4	5.2	11.6
	Male	≤30	79	6	3	10	5	66	0	2.5	2.5	7.4
		30–40	457	11	3	11	10	64	0	2	2.4	6.8
		40–50	543	10	2	8	14	71	0	2	3.4	11.7
		50–60	464	11	4	9	18	70	0	2	3.8	12.8
		>60	126	13	6	9	17	70	0	2	3.8	12.7
		All	1669	11	3	9	14	68	0	2	3.2	10.3
	All		3116	14	5	11	14	62	0	3	4.1	11.0
All	Female		12687	20	11	19	17	42	2	9	9.3	15.9
	Male		8921	14	6	13	17	57	0	5	6.1	14.0
All			21608	17	9	16	17	48	1	7	8.0	15.3

BC, blue collar; BMI, body mass index; C, clerical; P/M, professional/manager.

Table 4 Sickness absence (SA) by the HRA categories in different occupational groups by gender: means and the ratio of means

Interpretation of the HRA	Male				Female			
	SA days (N; mean; SD)			Ratio of means (95% CI)	SA days (N; mean; SD)			Ratio of means (95% CI)
Blue-collar workers								
No findings	274	5.18	10.50	Ref	198	6.95	17.83	Ref
Some symptoms	1060	7.18	15.73	**1.39 (1.07 to 1.87)**	1471	9.31	20.24	1.34 (0.96 to 2.11)
Health risk	691	8.33	16.56	**1.61 (1.23 to 2.19)**	843	12.04	22.42	**1.73 (1.24 to 2.74)**
WD risk	672	19.88	41.05	**3.84 (2.91 to 5.23)**	1394	20.64	38.19	**2.97 (2.15 to 4.67)**
Clerical employees								
No findings	532	2.43	12.20	Ref	425	4.48	16.6	Ref
Some symptoms	2198	3.10	10.61	1.28 (0.87 to 2.26)	3390	5.18	13.49	1.15 (0.84 to 1.80)
Health risk	953	4.08	11.57	**1.68 (1.11 to 3.00)**	1344	7.87	18.11	**1.75 (1.26 to 2.75)**
WD risk	872	9.80	25.2	**4.04 (2.69 to 7.19)**	2175	12.27	26.43	**2.74 (1.99 to 4.26)**
Professionals/managers								
No findings	232	1.77	4.71	Ref	95	2.00	3.76	Ref
Some symptoms	913	2.86	11.39	**1.61 (1.06 to 2.60)**	774	4.15	12.73	**2.08 (1.38 to 3.48)**
Health risk	292	3.89	12.15	**2.20 (1.29 to 3.69)**	220	5.98	14.81	**2.99 (1.79 to 5.20)**
WD risk	232	5.40	17.07	**3.05 (1.68 to 5.23)**	358	7.82	18.14	**3.91 (2.55 to 6.59)**

Bold values denote statistical significance at the p < 0.05 level.
HRA, health risk appraisal; WD, work disability.

increase in SA from 1.1 to 1.7 extra days missed annually compared with normal-weight employees in the USA.[21] It remains to be seen whether obesity acts as an additive risk factor for SA besides health problems.

In the present study, we evaluated how the HRA results predict SA in respondents from various industries and occupations and assessed the potential additive roles of self-reported health problems, occupational well-being and obesity. Our hypothesis was that self-reported health problems predict future SA, irrespective of gender and occupational group and that obesity has an additive effect.

METHODS
Study design, ethics and setting
The study design is a retrospective analysis of prospectively collected register data. The questionnaire data and SA register data were collected from one nationwide occupational health (OH) service provider's registers. Data privacy was strictly followed.

The study setting is OHS. Most Finnish employees use OHS for all their primary healthcare needs. In 2015, approximately 2.10 million Finnish employees (95% of the total workforce[24]) were covered by OHS. Besides 1.2 million preventive health examinations, OHS also performed 4.8 million illness-related visits.[25] The Finnish public social insurance system includes all lawful residents of Finland and provides health services and social service benefits for all members of the scheme, administered by The Social Insurance Institution of Finland (KELA). The SA benefit programme provides coverage for lost income

due to medically certified sickness or disease from day 1 until the person can work again up to 52 weeks. After that, long-term benefits from the disability benefits system provide coverage for lost income.

Participants
The study participants were working-age Finnish residents, aged 18–68 years, who had completed the HRA (n=22 515). An invitation to the HRA had been sent to 33 990 employees, of which 11 475 had not responded (response rate 66%). We used HRA results collected as a part of preventive OHS in 2012–2015 and archival data of SA covering 2011–2016. The inclusion criteria was a completed HRA. Exclusion criteria were >150 SA days in the 12 months preceding the HRA (n=119), granted disability benefit (n=689), missing data concerning occupational group (n=79) and loss to follow-up (n=77). Some respondents were in several exclusion categories. Finally, we analysed the data from 21 608 employees.

Patient and public involvement
Patients were not involved in the design, recruitment or conduct of the study.

Measurements
Explanatory variables
The classified results of the HRA were used as the primary exposure variable. Other exposure variables included problems with occupational well-being and obesity. Gender, age, occupational group and SA days before the HRA were treated as confounding factors.

Table 5 Predicting the odds for any sickness absence (SA) (logistic model) and the duration of SA, if any (zero-truncated negative binomial (NB) part) in the negative binomial Hurdle model: crude analysis

OG	Gender	HRA result category	N	Logistic model (0 vs >0)		Zero-truncated NB (>0)	
				OR	95% CI	RR	95% CI
BC	Female	(Intercept)		1.08	(0.82 to 1.43)	10.28	(7.81 to 13.53)
		No findings	198	1.00	Ref	1.00	Ref
		Some symptoms	1471	**1.78**	**(1.32 to 2.40)**	1.07	(0.80 to 1.42)
		Health risk	843	**2.05**	**(1.49 to 2.80)**	**1.36**	**(1.01 to 1.82)**
		WD risk, 1 risk factor	840	**3.13**	**(2.27 to 4.33)**	**1.58**	**(1.18 to 2.12)**
		WD risk, 2 risk factors	335	**4.91**	**(3.27 to 7.36)**	**2.36**	**(1.72 to 3.25)**
		WD risk, 3–5 risk factors	219	**5.59**	**(3.49 to 8.96)**	**3.43**	**(2.45 to 4.82)**
	Male	(Intercept)		1.23	(0.97 to 1.56)	6.64	(5.25 to 8.41)
		No findings	274	1.00	Ref	1.00	Ref
		Some symptoms	1060	**1.37**	**(1.04 to 1.79)**	1.27	(0.98 to 1.63)
		Health risk	691	**1.44**	**(1.08 to 1.91)**	**1.47**	**(1.13 to 1.92)**
		WD risk, 1 risk factor	456	**2.04**	**(1.49 to 2.79)**	**2.89**	**(2.19 to 3.82)**
		WD risk, 2 risk factors	147	**2.34**	**(1.51 to 3.63)**	**3.86**	**(2.70 to 5.51)**
		WD risk, 3–5 risk factors	69	**3.51**	**(1.83 to 6.71)**	**5.32**	**(3.41 to 8.30)**
C	Female	(Intercept)		0.62	(0.51 to 0.76)	8.05	(6.35 to 10.20)
		No findings	425	1.00	Ref	1.00	Ref
		Some symptoms	3390	**1.58**	**(1.28 to 1.94)**	0.87	(0.68 to 1.11)
		Health risk	1344	**1.94**	**(1.55 to 2.42)**	1.28	(0.99 to 1.65)
		WD risk, 1 risk factor	1403	**2.81**	**(2.24 to 3.51)**	**1.48**	**(1.15 to 1.91)**
		WD risk, 2 risk factors	500	**3.32**	**(2.54 to 4.36)**	**1.73**	**(1.30 to 2.30)**
		WD risk, 3–5 risk factors	272	**3.54**	**(2.56 to 4.88)**	**3.05**	**(2.21 to 4.21)**
	Male	(Intercept)		0.34	(0.28 to 0.42)	5.96	(4.51 to 7.86)
		No findings	532	1.00	Ref	1.00	Ref
		Some symptoms	2198	**1.36**	**(1.10 to 1.69)**	1.03	(0.77 to 1.37)
		Health risk	953	**1.46**	**(1.15 to 1.85)**	1.36	(0.99 to 1.86)
		WD risk, 1 risk factor	590	**2.59**	**(2.01 to 3.34)**	**1.84**	**(1.34 to 2.54)**
		WD risk, 2 risk factors	197	**3.19**	**(2.27 to 4.49)**	**3.22**	**(2.15 to 4.82)**
		WD risk, 3–5 risk factors	85	**4.16**	**(2.59 to 6.68)**	**3.83**	**(2.29 to 6.41)**
P/M	Female	(Intercept)		0.58	(0.38 to 0.88)	2.77	(1.59 to 4.83)
		No findings	95	1.00	Ref	1.00	Ref
		Some symptoms	774	1.23	(0.79 to 1.92)	**2.18**	**(1.26 to 3.78)**
		Health risk	220	1.45	(0.89 to 2.38)	**3.03**	**(1.65 to 5.58)**
		WD risk, 1 risk factor	254	**1.71**	**(1.06 to 2.78)**	**3.09**	**(1.71 to 5.60)**
		WD risk, 2 risk factors	69	**2.36**	**(1.25 to 4.46)**	**4.31**	**(2.09 to 8.90)**
		WD risk, 3–5 risk factors	35	**2.57**	**(1.16 to 5.69)**	**5.00**	**(2.10 to 11.9)**
	Male	(Intercept)		0.39	(0.29 to 0.52)	2.95	(1.80 to 4.81)
		No findings	232	1.00	Ref	1.00	Ref
		Some symptoms	913	1.06	(0.77 to 1.46)	**1.76**	**(1.12 to 2.78)**
		Health risk	292	1.38	(0.95 to 2.01)	**2.08**	**(1.23 to 3.51)**
		WD risk, 1 risk factor	176	1.43	(0.94 to 2.18)	**3.18**	**(1.76 to 5.73)**
		WD risk, 2 risk factors	36	**2.06**	**(1.00 to 4.21)**	1.35	(0.54 to 3.28)
		WD risk, 3–5 risk factors	20	**2.57**	**(1.02 to 6.46)**	2.51	(0.80 to 7.82)

Continued

Table 5	Continued						
				Logistic model (0 vs >0)		**Zero-truncated NB (>0)**	
OG	Gender	HRA result category	N	OR	95% CI	RR	95% CI

Logistic model refers to the model component for predicting membership to the subpopulation A with high propensity to zero absence, and zero-truncated NB to the component predicting the days on sick leave among the susceptible subpopulation B. To facilitate interpretation, for the zero-inflation part we have shown the ORs associated with the complementary propensity to having any sickness absence, that is, inclusion in subpopulation B.
Bold values denote statistical significance at the p < 0.05 level.
BC, blue collar; C, clerical; HRA, health risk appraisal; OG, occupational group; P/M, professional/manager; RR, risk ratio; WD, work disability.

The HRA result categories in declining priority order are 1) WD risk, 2) health risk, 3) some symptoms, 4) lifestyle issues and 5) no findings (table 1). Within the category 'WD risk', the results were further subdivided by the number of risk factors (1–5).

We constructed a dichotomous variable 'problems with occupational well-being', based on a series of questions modified from the JDCS model, the ERI model, the WLC theories and the presence of constant stress or dissatisfaction (table 2).[17–19] If any of the criteria were met, the respondent was classified as having a problem with occupational well-being.

BMI was categorised as underweight (<18.5), normal weight (18.5–25.0), overweight (25.0–30.0) and obese (>30.0). Normal weight was chosen as the reference class in the statistical models. Among males, underweight was combined with normal weight due to small numbers. Age was categorised into five classes: <30, 30–40, 40–50, 50–60 and >60 years. The age group 30–40 years was chosen the reference class in the statistical models. Occupational group was defined as blue-collar workers, clerical employees and professionals/managers. The number of SA days 12 months prior to the questionnaire was included as a continuous variable.

Sickness absence

We used SA days as the outcome variable. SA was operationalised as the accumulated number of days on sick leave during the 12-month follow-up after the survey. It includes the number of days and periods absent because of sickness. Overlapping and consecutive SA were combined. Maternity/paternity leave and absence from work to care for a sick child are not included in the SA.

The employer records the sick leave periods and dates when each SA starts and ends. If the SA is prescribed by the OH physician, the employer supplies the information to the OH care. In most cases, permanent employees are paid a full salary during their SA from the first day, up to 3 months. The employer receives sickness allowance from KELA after 10 working days, Sundays and other national holidays are excluded. The employee needs a medical certificate to qualify for sickness allowance. Mostly the blue-collar employees cannot complete their own certificates for any SA, while professional and manager employees must provide a written explanation for short SA and a medical certification for SA longer than 3 days. An employee may receive sickness allowance from KELA

for <1 year of WD due to the same illness. If an employee is unfit for work because of an illness for longer than a year, it is possible to claim a disability benefit. The evaluation of eligibility for WD benefits is transferred to the pension insurance companies if the illness lasts longer than 1 year.

Statistical methods

We analysed how the number of future SA days vary based on the results of the HRA, taking into account other exposure variables and potential confounding factors. Missing values in the questionnaire-based variables were imputed with the multiple imputation method MICE software with predictive mean matching. We are not aware of any systematic reasons or motives that would cause the non-response to be different in the HRA response categories. Based on our best knowledge, missing questionnaire data are missing at random. The following items were used as determinants when conducting the missing data imputation: gender, age, problems with occupational well-being, stress and fatigue, job satisfaction, BMI, all-cause SA both 12 months prior to and after and the lifestyle questions in the HRA (alcohol consumption, exercise, smoking).

There were complex interactions between gender and other variables in our data and we performed all analyses stratified by gender and occupational group as has been suggested earlier.[26]

Baseline characteristics are presented using descriptive statistics. Patterns with SA means and SD and the ratio of means with 95% CIs were calculated by the HRA result classes separately by gender and occupational group.

When modelling SA data, a special challenge is that a large number of employees have no absenteeism due to sickness.[6] Ordinary count data methods like Poisson or negative binomial regression models are not directly suitable for the analysis in case of the excess count of zero days. Our approach was to try mixture regression, zero-inflated negative binomial regression and the Hurdle model. The first two approaches yielded problems when estimating the model's parameters. We chose the Hurdle model, which provides a combination of the two statistical models: a binary model determines whether the outcome is zero or positive (logistic regression) and a truncated at zero count model for the positive part of the count data. We used the truncated negative binomial regression model because it accounts for the overdispersion present in count data. The estimated ORs (with 95% CI) based

on the binary part are reported in tabulated form. In the zero-truncated negative binomial part, the estimated risk ratios (RR) based on the regression coefficients of the HRA categories and covariates are reported with 95% CIs. The statistical analyses were performed using R V.3.4.4 software.

RESULTS

The average age of the participants was 45.3 years (SD 11; range 19–68) and 59% (n=12 687) were female, 6603 (32%) were blue-collar workers, 11 889 (56%) were clerical employees and 3116 (15%) belonged to the professional or manager category (figure 1). The non-respondents were slightly younger (average age 44.2 years, SD 12; t=−7.3; p<0.0001) than the respondents on the average. Also, males were less likely to respond than females with response rates 60% and 71%, respectively (χ^2=425.5; p<0.0001). The response rates were almost identical among blue-collar workers (65%), clerical employees (67%) and experts/managers (66%) (χ^2=14.3; p=0.0007).

A total of 172 331 days of SA were recorded in the study population during the 12-month follow-up. The distribution was heavily right-skewed in all age groups. Moreover, 48% had not been on SA at all, indicating a substantial zero component in the response distribution (table 3). The proportion of respondents with zero SA days was 35% in blue-collar males and 29% in females, 65% in clerical males and 46% in females, and 68% in professional/manager males and 55% in females.

The mean numbers of SA days among those with any SA were 18.2, 13.8, 11.0 days in blue collars, clericals and professionals/managers, respectively. An increasing trend of SA by age was observed among those with any SA. Females tend to have more SA days than males in all occupational groups. Twenty-six per cent of subjects reported 'WD risk factors', but their share of the total number of SA was 47% (table 3).

Belonging to the HRA category 'health risk' or 'WD risk factors' predicted higher mean values of SA during the follow-up, regardless of the occupational group or gender. The ratio of means of SA days varied between 2.7 and 4.0 among those with 'WD risk factors' and the reference category with no findings in the HRA, depending on gender and occupational group. The lower limit of the 95% CI was at the lowest 2.0. (table 4).

The results from fitting the Hurdle model are displayed in the table 5 (unadjusted model) and table 6 (adjusted model). The analyses were performed stratified by gender and occupational group due to complex interactions. We included the result of the HRA (six categories for blue collars and clericals and five categories for professionals/managers) as a covariate. The categories 'lifestyle issues and no findings" were combined as the reference class and the number of WD risk factors was analysed separately when possible. We excluded weight problems from the HRA 'health risk category' in the fully

adjusted model. The average number of SA among the susceptible to any SA followed the pattern blue-collar workers>clerical employees>professionals/managers in both genders. There was some evidence of an overall decreasing trend in the susceptibility to SA by increasing age. In males in both clerical and professional and managerial positions, the number of SA days tended to increase by age, but not in females or blue-collar males. SA prior to the HRA predicted both susceptibility for SA and the number of SA days during the follow-up. The presence of 'WD risk factors' predicted susceptibility to and the mean number of days on SA in all occupational groups. In the unadjusted model, HRA category 'WD risk factors' predicted the probability of SA (OR at the lowest 1.7 with 95% CI at the lowest 1.1 by occupational groups) for both genders (table 5). HRA category 'some symptoms' (OR at the lowest 1.4 with 95% CI at the lowest 1.0) and 'health risk' (OR at the lowest 1.4 with 95% CI at the lowest 1.1) predicted the probability for SA in blue collars and clericals. When all covariates were included (table 6), the ORs and RRs decreased. The number of earlier SA days predicted the probability of SA (OR at the lowest 1.03 with 95% CI at the lowest 1.02 by occupational groups) in the adjusted model. The presence of multiple 'WD risk factors' increased both susceptibility and/or the number of SA days. Problems with well-being at work predicted SA in the professional and managerial group in both genders. Overweight and/or obesity predicted SA in all professional groups in both genders. All these effects are additive, that is, adjusted for each other within each stratum.

DISCUSSION

Self-reported health problems in the HRA—musculoskeletal problems, depressive symptoms, sleep problems, constant stress and feeling of exhaustion and doubts about work ability—predicted future SA in both genders, regardless of occupational group. Of note, the larger the number of these problems, labelled as 'WD risk factors', the higher were the odds for any SA and the larger the number of SA days, if any. In Finland, the two largest categories of the causes of SA and permanent WD are musculoskeletal disorders and mental and behavioural disorders.[27] Also, problems with sleep,[28] constant stress,[29] exhaustion[30] and attitudes towards work ability[6] have predicted SA in earlier studies. It seems that using a questionnaire for self-rating of symptoms of the common causes of SA is a valid way to identify individuals at risk of SA, as the HR:s were relatively high in our study. Obesity and earlier sick leave days also predicted future SA in an additive fashion.

The strengths of the study include the registry-based, prospectively collected extensive data from various industries. Recorded SA data have several advantages: the quality of the data in terms of coverage, accuracy and consistency over time is better to that achievable via self-reports.[31] We were also able to control key potential

Table 6 Predicting the odds for any sickness absence (SA) (logistic model) and the duration of SA, if any (zero-truncated negative binomial (NB) part) in the negative binomial Hurdle model: all covariates included

OG	Gender	Explanatory variable	N	Logistic model (0 vs >0) OR	95% CI	Zero-truncated NB (>0) RR	95% CI
BC	Female	(Intercept)		1.13	(0.82 to 1.56)	8.48	(6.48 to 11.10)
		No findings	198	1.00	Ref.	1.00	Ref.
		Some symptoms	1471	**1.70**	**(1.25 to 2.31)**	0.97	(0.74 to 1.27)
		Health risk	843	**1.70**	**(1.23 to 2.37)**	1.05	(0.79 to 1.40)
		WD risk, 1 risk factor	840	**2.68**	**(1.91 to 3.75)**	1.21	(0.91 to 1.61)
		WD risk, 2 risk factors	335	**3.65**	**(2.37 to 5.61)**	**1.51**	**(1.11 to 2.06)**
		WD risk, 3–5 risk factors	219	**3.46**	**(2.08 to 5.75)**	**1.95**	**(1.39 to 2.74)**
		BMI>18.5 and ≤25	1866	1.00	Ref.	1.00	Ref.
		BMI≤18.5	53	1.13	(0.58 to 2.21)	1.42	(0.94 to 2.17)
		BMI>25 and ≤30	1210	1.09	(0.92 to 1.29)	**1.23**	**(1.10 to 1.39)**
		BMI>30	684	**1.27**	**(1.02 to 1.58)**	**1.29**	**(1.12 to 1.49)**
		Age>30 and ≤40	829	1.00	Ref.	1.00	Ref.
		Age≤30	701	0.84	(0.66 to 1.07)	1.06	(0.91 to 1.24)
		Age>40 and ≤50	1020	**0.66**	**(0.53 to 0.82)**	1.02	(0.88 to 1.17)
		Age>50 and ≤60	1067	**0.71**	**(0.57 to 0.88)**	1.08	(0.94 to 1.24)
		Age>60	289	**0.40**	**(0.30 to 0.54)**	1.02	(0.82 to 1.28)
		Problems in well-being	881	0.99	(0.82 to 1.20)	1.04	(0.92 to 1.17)
		Sick leaves before the questionnaire		**1.048**	**(1.039 to 1.057)**	**1.019**	**(1.016 to 1.022)**
	Male	(Intercept)		1.12	(0.83 to 1.51)	5.23	(4.05 to 6.75)
		No findings	274	1.00	Ref.	1.00	Ref.
		Some symptoms	1060	1.28	(0.97 to 1.69)	1.05	(0.82 to 1.34)
		Health risk	691	1.33	(0.98 to 1.79)	1.25	(0.97 to 1.61)
		WD risk, 1 risk factor	456	**1.63**	**(1.17 to 2.28)**	**2.06**	**(1.57 to 2.72)**
		WD risk, 2 risk factors	147	**1.70**	**(1.06 to 2.75)**	**2.16**	**(1.53 to 3.07)**
		WD risk, 3–5 risk factors	69	1.81	(0.89 to 3.68)	**3.29**	**(2.13 to 5.08)**
		BMI≤25	933	1.00	ref.	1.00	ref.
		BMI>25 and ≤30	1205	**1.22**	**(1.01 to 1.47)**	1.16	(1.00 to 1.35)
		BMI>30	527	**1.57**	**(1.22 to 2.02)**	1.15	(0.96 to 1.38)
		Age>30 and ≤40	702	1.00	Ref.	1.00	Ref.
		Age<30	475	0.88	(0.68 to 1.14)	1.04	(0.86 to 1.26)
		Age>40 and ≤50	684	**0.74**	**(0.59 to 0.94)**	**1.38**	**(1.16 to 1.64)**
		Age>50 and ≤60	659	**0.56**	**(0.44 to 0.71)**	1.16	(0.97 to 1.39)
		Age>60	177	**0.31**	**(0.22 to 0.44)**	1.32	(0.96 to 1.82)
		Problems in well-being	447	**1.28**	**(1.00 to 1.64)**	1.04	(0.87 to 1.23)
		Sick leaves before the questionnaire		**1.053**	**(1.042 to 1.065)**	**1.020**	**(1.015 to 1.024)**

Continued

Table 6 Continued

OG	Gender	Explanatory variable	N	Logistic model (0 vs >0) OR	Logistic model (0 vs >0) 95% CI	Zero-truncated NB (>0) RR	Zero-truncated NB (>0) 95% CI
C	Female	(Intercept)		0.67	(0.54 to 0.83)	6.27	(4.94 to 7.96)
		No findings	425	1.00	Ref.	1.00	Ref.
		Some symptoms	3390	**1.47**	**(1.19 to 1.82)**	0.90	(0.72 to 1.14)
		Health risk	1344	**1.72**	**(1.36 to 2.18)**	1.17	(0.91 to 1.51)
		WD risk, 1 risk factor	1403	**2.30**	**(1.82 to 2.91)**	**1.34**	**(1.05 to 1.71)**
		WD risk, 2 risk factors	500	**2.47**	**(1.85 to 3.29)**	1.27	(0.96 to 1.68)
		WD risk, 3–5 risk factors	272	**2.45**	**(1.73 to 3.47)**	**2.20**	**(1.60 to 3.03)**
		BMI>18.5 and ≤25	3775	1.00	Ref.	1.00	Ref.
		BMI≤18.5	81	0.82	(0.52 to 1.30)	0.68	(0.43 to 1.07)
		BMI>25 and ≤30	2213	1.08	(0.96 to 1.20)	1.10	(0.99 to 1.22)
		BMI>30	1192	**1.37**	**(1.18 to 1.59)**	**1.23**	**(1.08 to 1.4)**
		Age>30 and ≤40	1543	1.00	Ref.	1.00	Ref.
		Age<30	577	0.94	(0.77 to 1.14)	1.01	(0.84 to 1.21)
		Age>40 and ≤50	2036	**0.80**	**(0.70 to 0.93)**	1.07	(0.94 to 1.21)
		Age>50 and ≤60	2482	**0.63**	**(0.55 to 0.72)**	1.12	(0.99 to 1.27)
		Age>60	696	**0.45**	**(0.37 to 0.55)**	1.13	(0.93 to 1.36)
		Problems in well-being	1308	1.10	(0.96 to 1.25)	1.05	(0.94 to 1.19)
		Sick leaves before the questionnaire		**1.053**	**(1.046 to 1.060)**	**1.019**	**(1.016 to 1.022)**
	Male	(Intercept)		0.36	(0.29 to 0.45)	4.30	(3.16 to 5.85)
		No findings	532	1.00	Ref.	1.00	Ref.
		Some symptoms	2198	**1.41**	**(1.13 to 1.76)**	1.01	(0.76 to 1.33)
		Health risk	953	**1.50**	**(1.16 to 1.93)**	1.07	(0.78 to 1.46)
		WD risk, 1 risk factor	590	**2.50**	**(1.91 to 3.27)**	**1.46**	**(1.06 to 2.02)**
		WD risk, 2 risk factors	197	**2.70**	**(1.87 to 3.91)**	**2.42**	**(1.61 to 3.65)**
		WD risk, 3–5 risk factors	85	**3.51**	**(2.11 to 5.86)**	**2.85**	**(1.68 to 4.85)**
		BMI≤25	1687	1.00	Ref.	1.00	Ref.
		BMI>25 and ≤30	211	1.12	(0.97 to 1.29)	1.18	(0.99 to 1.39)
		BMI>30	734	**1.23**	**(1.01 to 1.49)**	**1.35**	**(1.08 to 1.69)**
		Age>30 and ≤40	1143	1.00	Ref.	1.00	Ref.
		Age<30	391	0.82	(0.64 to 1.05)	1.25	(0.93 to 1.69)
		Age>40 and ≤50	1179	**0.75**	**(0.63 to 0.90)**	**1.33**	**(1.08 to 1.62)**
		Age>50 and ≤60	1347	**0.64**	**(0.54 to 0.76)**	**1.40**	**(1.13 to 1.73)**
		Age>60	495	**0.47**	**(0.37 to 0.60)**	**1.80**	**(1.34 to 2.43)**
		Problems in well-being	551	1.18	(0.97 to 1.44)	0.95	(0.76 to 1.18)
		Sick leaves before the questionnaire		**1.040**	**(1.032 to 1.048)**	**1.016**	**(1.010 to 1.022)**

Continued

Table 6 Continued

OG	Gender	Explanatory variable	N	Logistic model (0 vs >0)		Zero-truncated NB (>0)	
				OR	95% CI	RR	95% CI
P/M	Female	(Intercept)		0.57	(0.36 to 0.91)	2.66	(1.52 to 4.65)
		No findings	95	1.00	Ref.	1.00	Ref.
		Some symptoms	774	1.12	(0.71 to 1.76)	**1.88**	**(1.10 to 3.20)**
		Health risk	220	1.17	(0.69 to 1.99)	**2.00**	**(1.09 to 3.67)**
		WD risk, 1 risk factor	254	1.43	(0.86 to 2.36)	**2.06**	**(1.15 to 3.68)**
		WD risk, 2 risk factors	69	1.69	(0.86 to 3.33)	**2.32**	**(1.12 to 4.82)**
		WD risk, 3–5 risk factors	35	2.02	(0.87 to 4.73)	**2.99**	**(1.25 to 7.13)**
		BMI>18.5 and ≤25	817	1.00	Ref.	1.00	Ref.
		BMI≤18.5	11	0.31	(0.06 to 1.45)	0.70	(0.08 to 5.83)
		BMI>25 and ≤30	416	**1.46**	**(1.14 to 1.87)**	1.18	(0.89 to 1.57)
		BMI>30	193	1.27	(0.90 to 1.79)	1.16	(0.80 to 1.67)
		Age>30 and ≤40	349	1.00	Ref.	1.00	Ref.
		Age<30	6	0.85	(0.48 to 1.50)	0.64	(0.33 to 1.22)
		Age>40 and ≤50	485	0.88	(0.66 to 1.17)	0.95	(0.70 to 1.31)
		Age>50 and ≤60	414	0.92	(0.68 to 1.24)	1.13	(0.81 to 1.56)
		Age>60	139	**0.63**	**(0.42 to 0.96)**	1.48	(0.91 to 2.39)
		Problems in well-being	194	0.95	(0.68 to 1.32)	**1.64**	**(1.13 to 2.38)**
		Sick leaves before the questionnaire		**1.040**	**(1.025 to 1.056)**	**1.027**	**(1.015 to 1.039)**
	Male	(Intercept)		0.41	(0.29 to 0.57)	2.47	(1.54 to 3.97)
		No findings	232	1.00	Ref.	1.00	Ref.
		Some symptoms	913	1.07	(0.77 to 1.49)	1.17	(0.76 to 1.82)
		Health risk	292	1.37	(0.91 to 2.05)	1.08	(0.63 to 1.86)
		WD risk, 1 risk factor	176	1.38	(0.89 to 2.14)	1.46	(0.82 to 2.62)
		WD risk, 2 risk factors	36	1.96	(0.92 to 4.17)	0.67	(0.27 to 1.66)
		WD risk, 3–5 risk factors	20	2.39	(0.92 to 6.18)	1.23	(0.41 to 3.75)
		BMI≤25	631	1.00	Ref.	1.00	Ref.
		BMI>25 and ≤30	805	1.11	(0.88 to 1.40)	1.04	(0.76 to 1.42)
		BMI>30	228	**1.53**	**(1.09 to 2.15)**	**1.61**	**(1.03 to 2.53)**
		Age>30 and ≤40	457	1.00	Ref.	1.00	Ref.
		Age<30	79	1.02	(0.62 to 1.70)	1.14	(0.60 to 2.16)
		Age>40 and ≤50	543	**0.71**	**(0.54 to 0.93)**	**1.74**	**(1.23 to 2.47)**
		Age>50 and ≤60	464	**0.68**	**(0.51 to 0.91)**	**2.11**	**(1.45 to 3.08)**
		Age>60	126	**0.64**	**(0.41 to 1.00)**	**2.64**	**(1.46 to 4.76)**
		Problems in well-being	154	1.05	(0.73 to 1.53)	**1.86**	**(1.15 to 3.01)**
		Sick leaves before the questionnaire		**1.032**	**(1.017 to 1.046)**	**1.023**	**(1.009 to 1.037)**

Logistic model refers to the model component for predicting membership to the subpopulation A with high propensity to zero absence, and zero-truncated NB to the component predicting the days on sick leave among the susceptible subpopulation B. To facilitate interpretation, for the zero-inflation part we have shown the ORs associated with the complementary propensity to having any sickness absence, that is, inclusion in subpopulation B.
Bold values denote statistical significance at the p < 0.05 level.
Health risk category does not include weight problems.
BC, blue collar; BMI, body mass index; C, clerical; OG, occupational group; P/M, professional/manager; RR, risk ratio; WD, work disability.

confounders, like age, gender, prior SA and occupational group. Another strength is that we used an HRA that has earlier shown to be able to identify employees with a high number of SA days. The results of the present study in a prospective setting in various industries were well in line with the earlier findings in the construction industry.[6 32–34] Earlier studies have provided evidence for the effectiveness of targeted health surveillance measures carried out based on HRA results.[4 33 34] The HRA used in the present study was able to identify a subgroup with WD risk in an earlier study.[6] A randomised trial was carried out within the above-mentioned high-risk group, half of which received an invitation to OH services for targeted health surveillance and half of them received usual care. The difference between the targeted intervention group and the control group was 10.6 days in favour of the intervention during the 12-month follow-up.[32] The total cost of healthcare was on average €180 per person less in the intervention group than in the usual care group.[33] Thus, the HRA used in the present study seems to focus on the essential health problems.[32]

We consider 'healthy worker effect' as a potential limitation of our study. It might be present if employees with worse health level had not responded or they are less likely to hire.[35] This potential bias would underestimate the associations as the respondents would be healthier, and possibly have less SA than non-respondents. Similar bias would potentially result from a 'healthy worker survival effect', which means that only healthiest and strongest will remain in the working life.[36] Moreover, we did not include those who were on long-term sick leave or those who had already been granted a disability benefit before the HRA. All this might underestimate the associations. It may also be possible that the healthiest employees might not respond to the HRA, which would have an opposite effect on our estimates. In our study, the participants were slightly older than non-participants and participation rate was higher among females than males. This would potentially overestimate the associations if the respondents had more illnesses than non-respondents.

Analysis of the predictors and determinants of SA is difficult with traditional statistical methods because a substantial fraction is clustered at zero SA days. Also, the residual variability in the non-zero part of the SA distribution exceeds that predicted by a Poisson model for counts. Although the Hurdle model[37] was perhaps not able to deal with all the complexity associated with this type of response variable, among computationally feasible approaches it is clearly more appropriate than the simpler alternative models in dealing with both the extra-zero component and the overdispersion. However, the residual collinearity between age, weight, WD risk factors and problems with occupational well-being can cause imprecise estimates of the coefficient values and therefore the resulting out-of-sample predictions may be imprecise.

Our results provide further support to the fact that obesity is an independent risk factor for SA, which is in line with other recent studies.[20 22 38] A strong connection between prior SA and the later SA has been found earlier,[39 40] also in line with our findings. Blue-collar workers had the highest and the professional-level/manager-level employees had the lowest level of SA days in our study, the same way as in earlier studies.[41–44] We found that females tend to have more SA, as has been reported before.[45] Females also report more often symptoms or other health problems,[15] than males.

We conclude that the use of an HRA with predictive validity can improve the quality of health surveillance: screening questionnaires seem appropriate for targeting efforts to employees in need. They seem useful in identifying symptoms and signs that predict SA. These findings have implications for both management and the healthcare system in the prevention of WD. Further research is needed to assess whether the HRA also predicts long-term WD. Also, the effectiveness and cost-effectiveness of the targeted health surveillance for the risk groups warrant further research.

Author affiliations
[1]Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland
[2]Terveystalo, Helsinki, Finland
[3]Finnish Institute of Occupational Health, Tampere, Finland
[4]Faculty of Information Technology and Communication Sciences, Tampere University, Tampere, Finland
[5]Orthopedics and Traumatology, University of Helsinki, Helsinki, Finland
[6]Evalua International, Helsinki, Finland

Twitter Minna Pihlajamäki @k_pihlajamaki

ORCID iD
Simo Taimela http://orcid.org/0000-0001-6755-2983

REFERENCES

1 Aumayr-Pintar C, Moulaï K. European Foundation for the Improvement of Living and Working Conditions. In: *Developments in working life in Europe: EurWORK annual review 2015*. Luxembourg: Publications Office of the European Union, 2016. https://

digitalcommons.ilr.cornell.edu/cgi/viewcontent.cgi?article=1575&context=intl

2 Laaksonen M, Kääriä S-M, Leino-Arjas P, *et al*. Different domains of health functioning as predictors of sickness absence – a prospective cohort study. *Scand J Work Environ Health* 2011;37:213–8.

3 van Hoffen MFA, Joling CI, Heymans MW, *et al*. Mental health symptoms identify workers at risk of long-term sickness absence due to mental disorders: prospective cohort study with 2-year follow-up. *BMC Public Health* 2015;15:1235.

4 Stange B, McInerney J, Golden A, *et al*. Integrated approach to health screening of former department of energy workers detects both occupational and non-occupational illness. *Am J Ind Med* 2016;59:200–11.

5 Airaksinen J, Jokela M, Virtanen M, *et al*. Prediction of long-term absence due to sickness in employees: development and validation of a multifactorial risk score in two cohort studies. *Scand J Work Environ Health* 2018;44:274–82.

6 Taimela S, Läärä E, Malmivaara A, *et al*. Self-Reported health problems and sickness absence in different age groups predominantly engaged in physical work. *Occup Environ Med* 2007;64:739–46.

7 Szubert Z, Makowiec-Dąbrowska T, Merecz D, *et al*. Predictors of short- and long-term sickness absence in female post office workers in Poland. *Int J Occup Med Environ Health* 2016;29:539–62.

8 Loisel P, Durand MJ, Berthelette D, *et al*. Disability prevention: new paradigm for the management of occupational back pain. *Disease Management and Health Outcomes* 2001;9:351–60.

9 Momsen A-MH, Stapelfeldt CM, Nielsen CV, *et al*. Screening instruments for predicting return to work in long-term sickness absence. *Occup Med* 2017;67:101–8.

10 Kinnunen U, Nätti J. Work ability score and future work ability as predictors of register-based disability pension and long-term sickness absence: a three-year follow-up study. *Scand J Public Health* 2018;46:321–30.

11 Lundin A, Kjellberg K, Leijon O, *et al*. The association between Self-Assessed future work ability and long-term sickness absence, disability pension and unemployment in a general working population: a 7-year follow-up study. *J Occup Rehabil* 2016;26:195–203.

12 Clausen T, Burr H, Borg V. Do psychosocial job demands and job resources predict long-term sickness absence? an analysis of register-based outcomes using pooled data on 39,408 individuals in four occupational groups. *Int Arch Occup Environ Health* 2014;87:909–17.

13 Spanier K, Peters E, Michel E, *et al*. Associations between organizational injustice and work ability, self-reported disability days, and medical consultations: cross-sectional findings from employees with prior sickness absence payments. *Int Arch Occup Environ Health* 2017;90:789–97.

14 Vahtera J, Pentti J, Kivimäki M. Sickness absence as a predictor of mortality among male and female employees. *J Epidemiol Community Health* 2004;58:321–6.

15 Kivimäki M, Head J, Ferrie JE, *et al*. Sickness absence as a global measure of health: evidence from mortality in the Whitehall II prospective cohort study. *BMJ* 2003;327.

16 Knardahl S, Johannessen HA, Sterud T, *et al*. The contribution from psychological, social, and organizational work factors to risk of disability retirement: a systematic review with meta-analyses. *BMC Public Health* 2017;17:176–31.

17 Kinman G, Jones F. Effort-reward imbalance, over-commitment and work-life conflict: testing an expanded model. *J Manag Psychol* 2008;23:236–51.

18 Harter Griep R, Rotenberg L, Chor D, *et al*. Beyond simple approaches to studying the association between work characteristics and absenteeism: combining the DCs and eri models. *Work Stress* 2010;24:179–95.

19 Yu S, Lu M-L, Gu G, *et al*. Association between psychosocial job characteristics and sickness absence due to low back symptoms using combined DCs and eri models. *Work* 2015;51:411–21.

20 Janssens H, Clays E, Kittel F, *et al*. The association between body mass index class, sickness absence, and presenteeism. *J Occup Environ Med* 2012;54:604–9.

21 Andreyeva T, Luedicke J, Wang YC. State-level estimates of obesity-attributable costs of absenteeism. *J Occup Environ Med* 2014;56:1120–7.

22 Roos E, Laaksonen M, Rahkonen O, *et al*. Weight change and sickness absence--a prospective study among middle-aged employees. *Eur J Public Health* 2015;25:263–7.

23 Lehnert T, Sonntag D, Konnopka A, *et al*. The long-term cost-effectiveness of obesity prevention interventions: systematic literature review. *Obes Rev* 2012;13:537–53.

24 Lappalainen K, Aminoff M, Hakulinen H, *et al*. Työterveyshuolto Suomessa vuonna 2015: ja kehitystrendi 2000 - 2015. Tietoa työstä 2016. Originally published in Finnish. [Occupational Health Care in Finland 2015: and the Development Trend 2000 - 2015]. Published on the internet. Available: https://www.julkari.fi/handle/10024/131293 [Accessed May 2018].

25 Social Insurance Institution of Finland Kela. Kelan työterveyshuoltotilasto 2015. Originally published in Finnish. [Social Insurance Institution of Finland: Official statistics of Finland]. Published on the internet in June, 2017. Available: http://hdl.handle.net/10138/212975 [Accessed May 2018].

26 Messing K, Punnett L, Bond M, *et al*. Be the fairest of them all: challenges and recommendations for the treatment of gender in occupational health research. *Am J Ind Med* 2003;43:618–29.

27 Kaila-Kangas L, Haukka E, Miranda H, *et al*. Common mental and musculoskeletal disorders as predictors of disability retirement among finns. *J Affect Disord* 2014;165:38–44.

28 Lallukka T, Kaikkonen R, Härkänen T, *et al*. Sleep and sickness absence: a nationally representative register-based follow-up study. *Sleep* 2014;37:1413–25.

29 Virtanen M, Vahtera J, Pentti J, *et al*. Job strain and psychologic distress influence on sickness absence among Finnish employees. *Am J Prev Med* 2007;33:182–7.

30 Suominen S, Vahtera J, Korkeila K, *et al*. Job strain, life events, and sickness absence: a longitudinal cohort study in a random population sample. *J Occup Environ Med* 2007;49:990–6.

31 Stapelfeldt CM, Jensen C, Andersen NT, *et al*. Validation of sick leave measures: self-reported sick leave and sickness benefit data from a Danish national register compared to multiple workplace-registered sick leave spells in a Danish municipality. *BMC Public Health* 2012;12:661.

32 Taimela S, Malmivaara A, Justén S, *et al*. The effectiveness of two occupational health intervention programmes in reducing sickness absence among employees at risk. two randomised controlled trials. *Occup Environ Med* 2008;65:236–41.

33 Taimela S, Justén S, Aronen P, *et al*. An occupational health intervention programme for workers at high risk for sickness absence. cost effectiveness analysis based on a randomised controlled trial. *Occup Environ Med* 2008;65:242–8.

34 Taimela S, Aronen P, Malmivaara A, *et al*. Effectiveness of a targeted occupational health intervention in workers with high risk of sickness absence: baseline characteristics and adherence as effect modifying factors in a randomized controlled trial. *J Occup Rehabil* 2010;20:14–20.

35 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.

36 Nordström K, Hemmingsson T, Ekberg K, *et al*. Sickness absence in workplaces: does it reflect a healthy hire effect? *Int J Occup Med Environ Health* 2016;29:315–30.

37 Smithson M, Merkle EC. *Generalized linear models for categorical and continuous limited dependent variables. Taylor & Francis group, Boca Raton, FL 33487-2742*. CRC Press, 2013: 127–32.

38 Neovius K, Johansson K, Kark M, *et al*. Obesity status and sick leave: a systematic review. *Obes Rev* 2009;10:17–27.

39 Cohen A, Golan R. Predicting absenteeism and turnover intentions by past absenteeism and work attitudes. *Career Development International* 2007;12:416–32.

40 Laaksonen M, He L, Pitkäniemi J. The durations of past sickness absences predict future absence episodes. *J Occup Environ Med* 2013;55:87–92.

41 Benavides FG, Benach J, Mira M, *et al*. Occupational categories and sickness absence certified as attributable to common diseases. *Eur J Public Health* 2003;13:51–5.

42 Piha K, Laaksonen M, Martikainen P, *et al*. Interrelationships between education, occupational class, income and sickness absence. *Eur J Public Health* 2010;20:276–80.

43 Melchior M, Krieger N, Kawachi I, *et al*. Work factors and occupational class disparities in sickness absence: findings from the GAZEL cohort study. *Am J Public Health* 2005;95:1206–12.

44 Pekkala J, Blomgren J, Pietiläinen O, *et al*. Occupational class differences in long sickness absence: a register-based study of 2.1 million Finnish women and men in 1996-2013. *BMJ Open* 2017;7:e014325.

45 Laaksonen M, Mastekaasa A, Martikainen P, *et al*. Gender differences in sickness absence--the contribution of occupation and workplace. *Scand J Work Environ Health* 2010;36:394–403.