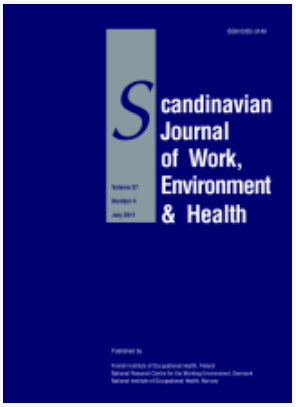

Original article

Scand J Work Environ Health 2013;39(1):57-65
doi:10.5271/sjweh.3298

Does perceived work ability improve after a multidisciplinary preventive program in a population with no severe medical problems – the Finnish Public Sector Study

by Saltychev M, Laimi K, Oksanen T, Pentti J, Kivimäki M, Vahtera J

Affiliation: Department of Rehabilitation, Turku University Hospital, PO Box 52, FI-20521 Turku, Finland. mikhail.saltychev@gmail.com

Refers to the following texts of the Journal: 2010;36(5):404-412 2011;37(1):1-5 2012;38(1):27-37

Key terms: Finland; Finnish Public Sector Study; multidisciplinary preventive program; primary prevention; propensity score; vocational rehabilitation; work ability; work ability index; work disability

Original article

Scand J Work Environ Health. 2013;39(1):57–65. doi:10.5271/sjweh.3298

Does perceived work ability improve after a multidisciplinary preventive program in a population with no severe medical problems? The Finnish Public Sector Study

by Mikhail Saltychev, MD,[1] Katri Laimi, PhD,[2, 3] Tuula Oksanen, PhD,[4] Jaana Pentti, BSc,[4] Mika Kivimäki, PhD,[4, 5] Jussi Vahtera, PhD [2, 4]

Saltychev M, Laimi K, Oksanen T, Pentti J, Kivimäki M, Vahtera J. Does perceived work ability improve after a multidisciplinary preventive program in a population with no severe medical problems? The Finnish Public Sector Study. *Scand J Work Environ Health.* 2013;39(1):57–65. doi:10.5271/sjweh.3298

Objectives This study examines the short- and long-term effects of a multidisciplinary preventive program on perceived work ability in a population with no severe medical problems.

Methods Altogether 859 public sector employees who participated in the program in 1997–2005 and their 2426 propensity-score-matched controls were studied prospectively. Propensity scores for probability of being granted participation in the program were calculated based on the data on health, health-risk behaviors, and work-related characteristics that were gathered from repeat responses to a survey, national health registers, and employers' records. Mean scores of perceived work ability (PWA) and prevalence ratios (PR) of suboptimal PWA were calculated after a short-term (mean 1.7 years, up to 4.6 years) and a long-term (mean 5.8 years, up to 9.2 years) follow-up.

Results No beneficial effects were observed with respect to work ability. In comparison to controls, the participants' risk of suboptimal PWA was actually slightly higher after both the short- [PR 1.23, 95% confidence interval (95% CI) 1.10–1.39] and long-term (PR 1.18, 95% CI 1.06–1.31) follow-ups.

Conclusions These data suggest that the vocationally oriented multidisciplinary preventive program was ineffective in improving work ability among participants with no severe medical problems.

Key terms Finland; primary prevention; propensity score; vocational rehabilitation; work disability; work ability index.

Responsible for the loss of almost 10% of the entire gross domestic product (GDP) in OECD countries, early retirement is a growing threat to the economics of developed countries (1). Numerous preventive interventions, containing both medical and vocational methods to improve the work ability of workers, have been suggested for tackling this threat (2–4). These measures target either high- or low-risk populations. The difference between these two approaches can be understood as a difference between secondary and primary prevention. Secondary preventive methods are aimed at workers who already have chronic health problems and a work ability that has started to deteriorate. In a low-risk population, however, there is only a threat of work ability deteriorating, for example, due to work-related physical or psychological stress or minor symptoms of health problems, but otherwise the workers are relatively healthy.

Reducing the risk of early retirement due to back pain is the most common goal of previously reported studies, with return to work being their most common outcome (4). The role of psychological factors in the development of long-term work disability has been suggested as being even more important than the role of biomechanical or biomedical factors (3). Therefore,

1 Department of Rehabilitation, Turku University Hospital, Turku, Finland.
2 Department of Public Health, University of Turku and Turku University Hospital, Turku, Finland.
3 Department of Physical and Rehabilitation Medicine, Turku University Hospital, Turku, Finland.
4 Finnish Institute of Occupational Health, Helsinki and Turku, Finland.
5 Department of Behavioral Sciences, University of Helsinki, Helsinki, Finland.

Correspondence to: Mikhail Saltychev, Department of Rehabilitation, Turku University Hospital, PO Box 52, FI-20521 Turku, Finland. [E-mail: mikhail.saltychev@gmail.com]

the content of the interventions in question usually follows the multidisciplinary biopsychosocial model, even though for workers with chronic back pain, adding cognitive behavioral therapy to work conditioning programs did not show improvements in the effectiveness compared to single-professional physical programs (5). Probably the most studied type of such prevention is the work-conditioning program widely used in the United States (US) since the 1970s (6). While preventive interventions of work disability have been found to be potentially effective, there is no agreement on exactly how "early" the prevention should be applied. Multidisciplinary preventive programs have been reported to be effective in improving the work ability of persons who already have substantial health problems (2). Instead, individually based primary preventive interventions have been found to have only a limited effect, for example, on musculoskeletal and cardiovascular diseases (7, 8). However, it has been suggested that interventions implemented before the occurrence of severe symptoms of disease and impairment may be effective in preventing the development of persistent work disability (3, 4). Such interventions may have at least the following two potential benefits: (i) a direct preventive effect on the deterioration of work ability and (ii) an indirect influence by releasing limited resources for more effective, but also more expensive, secondary prevention. Limited evidence is available for the effectiveness of stress management training interventions to reduce work-related stress levels (9, 10).

We studied vocationally oriented multidisciplinary rehabilitation (VOMR), which is a Finnish primary prevention of work disability method. This program was invented in the early 1980s as a measure of secondary prevention, more or less resembling the purposes and structure of the work-conditioning program. However, since the later 1980s, it has become a primary form of prevention, targeting workers who do not have severe medical problems but may have experienced some minor symptoms caused by work-related strain. The theoretical base for this transition is the assumption that intervention is more efficient if it is applied at a very early stage of disease, before persistent incapacity for work forms. Existing data on the effectiveness of this program are limited and inconsistent (11, 12). Previously, we have found no beneficial effect of VOMR on the sickness absence or perceived health of rehabilitants (11, 12). Instead, the VOMR selection process is likely to fail to recognize workers with higher risks of work disability and favors those with lower risks (13). Although not beneficial in terms of general health, the program could be beneficial in terms of perceived work ability (PWA) as people with severe illnesses may still use their partial work capacity to work in their specific jobs (14). Because the prevention program is extremely costly (over 30 million euros annually) (15), widely used in Finland, and specifically tailored to improve work ability rather than general health, it is important to examine the effects of the program on PWA.

As the program is directed towards the long-term prevention of work disability, our study aimed to evaluate both the short- and long-term effects of VOMR with respect to improving PWA in a comparison made with propensity-score-matched controls who had the same likelihood of being selected for the intervention as the cases according to their baseline (pre-rehabilitation) characteristics.

The conventional matching procedure identifies the treated and untreated persons by matching them exactly, for example, according to demographic characteristics such as gender, age group, and socioeconomic position. After the matching, these characteristics are distributed evenly in the groups. This procedure enables the selection to be controlled for only certain variables of the countless number of observed and unobserved characteristics. A conventional exact-matching procedure may work sufficiently when the number of variables is relatively small. However, when there are many matching variables, the list of possible variables is large, the match is difficult to achieve for each observed variable (the multi-dimensional problem) (16), and the number of persons who remain unmatched increases. Propensity-score matching avoids the multi-dimensional problem as it transforms the matching process into a single dimensional one (16). The procedure of propensity-score matching is performed for a single calculated variable: the propensity score. The calculated propensity score describes the probability of each person being selected to the intervention by a large number of baseline characteristics. Propensity-score matching is a good choice for creating a control group in an epidemiologic study when true randomization is not possible (17). It is applicable to large study populations when data on the individuals' characteristics are gathered before the intervention (18, 19). In contrast to randomization or instrumental variables analysis, propensity-score matching cannot rule out unmeasured third factors although it provides control for all of the measured confounders (20).

Methods

Study population

This research was part of the Finnish Public Sector Study, an on-going prospective cohort study of employees working in 10 municipalities and 21 hospitals. The study comprises 151 618 employees with a ≥6-month job contract in any year between 1991–2005. Data

on psychosocial factors at work, individual factors, health, and health behaviors have been gathered from responses to questionnaire surveys. All of the participants have been linked to their employers' records and also to national health registers. The Ethics Committee of the Finnish Institute of Occupational Health (FIOH) approved the study.

We used data from repeat responses to the same survey made in 1997–1998, 2000–2002, 2004–2005, and 2008–2009. First, we included all 53 416 employees who had responded to baseline surveys either in 1997–1998 or 2000–2002 (70% of eligible participants responded, 81% of whom were women). We included only those who responded to two consecutive identifiable follow-up surveys in 2000–2002, 2004–2005, or 2008–2009. We excluded all employees who had been granted any rehabilitation before the baseline survey (N=4176) or had missing data for any of the matching variables (N=3856). The potential participants were anyone who had entered the VOMR program in 1997–2005 between the baseline and the 1st follow-up survey (N=1000). This process yielded a sample of 24 100 employees who were eligible for matching (887 cases and 23 213 controls). Using propensity-score matching, we identified the study population of 859 participants and 2426 non-cases of VOMR (figure 1).

Intervention – vocationally oriented multidisciplinary rehabilitation (VOMR)

VOMR is a multidisciplinary, early preventive program that targets workplaces and occupations in which workers are subjected to considerable physical, mental, or social strain that may easily lead to health problems and a deterioration of work capacity. Its primary aim is the prevention of work disability. The participants generally have only minor health problems, as the VOMR selection criteria include, among others, an absence of recent long-term sick leave or a severe illness decreasing work capacity or any indication of alcohol or drug abuse. In 2010, the median age of the employees participating in VOMR was 50 years (15). The participants' need for rehabilitation is recognized by occupational physicians, and each group of rehabilitants usually has the same employer and/or profession. The final acceptance of selected employees is determined in social insurance offices around the country. Rejection is relatively rare (<13% of all VOMR applications were rejected in 2010) (15).

At the time of the study in 1997–2005, VOMR was primarily implemented as an in-patient program that usually contained three or four periods of in-patient, extensive, multi-modal, and multi-professional rehabilitation (total: 15–21 days) implemented as group-based (8–10 persons) supervised activity 4–6 hours per day. The multi-professional team consisted of a physician, physiotherapist, psychologist, social worker, and vocational rehabilitation specialist. In addition, a nurse, occupational therapist, occupational physiotherapist, and nutritionist were often involved. The modalities included physiotherapy and physical and psychological education. All of the activities were directed towards improving the physical and mental health status of the participants, enhancing their stress management, and encouraging a healthy lifestyle (eg, improving dietary habits and leisure-time physical activity, reducing or quitting smoking and drinking). The concept of physical training included an individual assessment of the participants and a plan for exercising at home and during workday breaks as well as during the next in-patient period of the program. It also included ergonomic education and exercises performed in groups. Problems at worksites, such as work-related strain and ways to manage it, were discussed in group-based sessions with a psychologist, social worker, and physician. The program required representatives from the worksite (usually a supervisor and an occupational physician) to participate in joint, 1-day group sessions. Sometimes the physical work environment was adjusted. Although VOMR is implemented in different independent rehabilitation facilities, the Social Insurance Institution of Finland strictly defines the inclusion criteria, the structure of the program, the multi-professional team composition, the modalities, and the assessment tests. The program follows this pre-determined plan, but the content of the group-based sessions may differ slightly based on the occupational characteristics of the participants in the group. Between the in-patient periods, the participants are expected to follow an individual exercise program at home, which usually consists of self-reliant physical activities and psychological exercises.

The participants do not work during the in-patient periods, and the entire program is free of charge. The participants receive a so-called "rehabilitation compensation", which is about 75% of the participant's usual salary (a minimum of 22 euros per workday). The participants also receive compensation for travelling expenses. Employers are not financially compensated for hiring temporary agency workers or substitutes.

Outcome – perceived work ability

PWA was assessed using the first item of the Work Ability Index (WAI), developed by FIOH and widely used for >20 years. The index has been found to correlate with the physical and psychological condition of employees (21, 22). Derived from the (WAI), the 11-point scale has been found to be a reliable and easy tool; its validity is also comparable with that of the full WAI (21, 23, 24). The perceived work ability assessment was based on three repeat responses (at baseline, the end

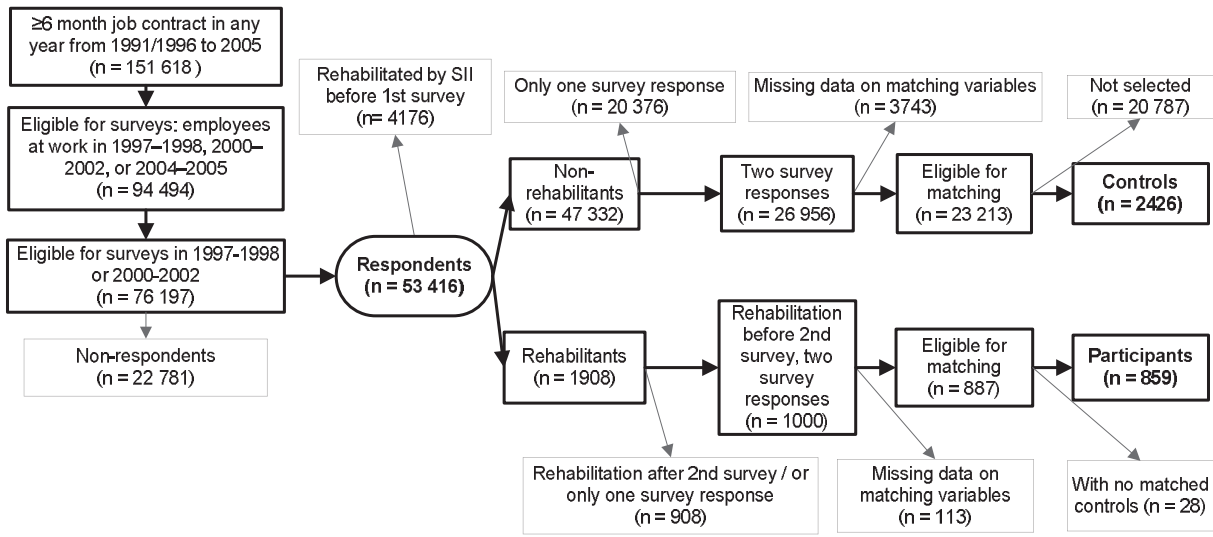

Figure 1. Case–control selection flowchart.

of the short-term follow-up 1.7 years later, and the end of the long-term follow-up 5.8 years later) to a standard single-item question concerning "current work ability compared with the lifetime best", with a possible score ranging from 0 ("completely unable to work") to 10 ("work ability at its best"). The score was dichotomized (0–7 versus 8–10) to differentiate between those with suboptimal and optimal PWA. To examine whether the results were dependent on the cut-off point used, we also used the PWA score as a continuous variable.

Statistical analysis

The control group was selected by the propensity-score matching to approximate the exchangeability of the comparison groups. The propensity score is the conditional probability of being assigned "treatment", in this case VOMR, given the observed covariates (18, 19). In other words, this approach ascertains – in theory – whether the cases and controls differ only in the receipt of VOMR. Using binary logistic regression models with "being granted VOMR" (dichotomous outcome) as the dependent variable, we calculated propensity scores for all 24 100 employees eligible for this study (see figure 1). We also included the baseline self-rated work ability score divided into tertiles (0–7, 8–9, and 10) and 24 other pre-treatment characteristics and their interactions with gender, socioeconomic status, and age group as dependent variables for all 96 terms in the model. [For the associations between the baseline characteristics and the subsequent receipt of VOMR, see Saltychev et al (11).] For each person, the modeling gives a score ranging from 0–1 (ie, his or her probability to be a case as a function of the predictor terms). Although the distribution of the propensity scores among the rehabilitants (range 0–0.64) and non-rehabilitants (range 0–0.58) was practically the same, a very low propensity score (0–0.03) was found for the majority (71%) of the non-rehabilitants, but for only a minority (26%) of the rehabilitants (figure 2). Once the propensity score was estimated, each case was matched with 1–3 controls (non-VOMR recipients) according to a predefined caliper width of 0.01, and the unmatched cases (N=28) were discarded. The balance achieved by the matching was studied using the Chi-square test. The case–control selection flow of the study is presented in figure 1.

Repeated-measures analysis

We applied repeated-measures generalized linear models using generalized estimating equations (GEE) for calculating the prevalence of suboptimal PWA at baseline and the follow-up measurements and the corresponding prevalence ratios (PR) and their 95% confidence intervals (95% CI) using log-binomial regression analysis (25, 26). Then we calculated the changes in the mean PWA score and their 95% CI using regression models for the continuous variables. In these models, we included the main effect of time (3 levels) and group (2 levels) and their interaction term "time × group". The interaction tested whether the PR or the mean difference between the cases and controls were the same at all three time points. The 1st follow-up survey (short-term follow-up) averaged 1.7 [standard deviation (SD) 1.02, range 0.003–4.55] years, and the average of the 2nd follow-up survey (long-term follow-up) was 5.8 (SD 1.14, range 3.01–9.16) years after the intervention.

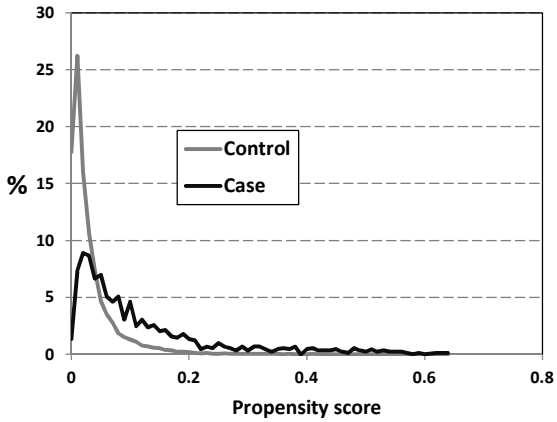

Figure 2. Distribution of propensity scores in the participants (case) and control group.

Sensitivity analyses

To examine the extent to which the results were sensitive to the method applied, we compared the results obtained by the propensity-score matching to those obtained by adjustment. We calculated the changes in the mean PWA score among the rehabilitants compared with the changes for all 23 213 eligible non-rehabilitants while adjusting either for age, gender, and occupational status (a conventional approach) or for the propensity score.

All of the statistical analyses were performed using SAS 9.2 software (SAS Institute Inc, Cary, NC, USA).

Results

Data on the baseline characteristics of the participants were gathered an average of 1.8 (SD 1.1) years before the intervention (table 1). The mean score of the self-reported baseline PWA was 8.15 for the VOMR participants and 8.17 for the controls (P=0.78). Compared with that of the controls, the PWA deteriorated more among the VOMR participants over time, as indicated by a significant time × group interaction for both the prevalence of suboptimal PWA [chi 2(df)=5.91, P=0.05] and the mean PWA score [chi 2(df)=6.89, P=0.03]. Although there was no difference in the prevalence of suboptimal PWA among the VOMR participants compared with that of the controls at baseline, suboptimal PWA among the participants was 1.23 times (95% CI 1.10–1.39) more likely at the short-term follow-up and 1.18 (95% CI 1.06–1.31) more likely at the long-term follow-up (table 2). Table 2 also shows the mean values of PWA over the follow-up. Compared with the baseline level, the mean PWA score decreased by 0.41 (95% CI 0.31–0.57) points among the VOMR participants but only by 0.26 (95% CI 0.20–0.32) points among the controls (P=0.005) at the short-term follow-up; at the long-term follow-up, these mean differences were 0.58 (95% CI 0.46–0.69) and 0.46 (95% CI 0.39–0.53), respectively.

As a sensitivity analysis, we examined how the results changed when we used both conventional and propensity score adjustment in a data-set composed of all 887 rehabilitants and 23 213 eligible non-rehabilitants as our analytical approach (figure 3). After adjustment for demographics, the mean PWA score had decreased by 0.42 (95% CI 0.33–0.51) points from the baseline level among the VOMR participants and by 0.29 (95% CI 0.27–0.31) points among the non-rehabilitants in the short-term follow-up; in the long-term follow-up, these mean differences were 0.58 (95% CI 0.46–0.69) and 0.53 (95% CI 0.51–0.55), respectively (time × group interaction P=0.03). After adjustment for the propensity score, the corresponding mean score differences were 0.42 (95% CI 0.33–0.51) and 0.29 (95% CI 0.27–0.31) in the short-term follow-up and 0.57 (95% CI 0.46–0.69) and 0.53 (95% CI 0.50–0.55) in the long-term follow-up, respectively (time × group interaction P=0.02). Thus, all three approaches gave practically the same results for the participants (the mean score level decreased by 0.41–0.42 points on average from the baseline level at the short-term follow-up and by 0.57–0.58 at the long-term follow-up). However, in relation to the controls, the decline in mean PWA was steeper when adjustments were used than when propensity-score matching was used (0.29 versus 0.26 at the short-term follow-up, and 0.53 versus 0.46 at the long-term follow-up, respectively).

Discussion

In this study during the follow-up of up to 9 years, PWA declined more among the 859 VOMR program participants than their 2426 propensity-score matched controls.

Most previous reports on the positive effect of multidisciplinary rehabilitation on work ability have evaluated interventions in high-risk populations, where, by definition, the need for rehabilitation is induced by functional impairment. Instead, due to the preventive character of its inclusion criteria, VOMR intervention is placed on a timeline before the occurrence of impairment or even disease.

One study from Germany reported improvement in the PWA of ageing bus drivers in a 1-year follow-up after a health promotion program that was slightly comparable to VOMR (27). In addition, two previous studies on the effectiveness of VOMR have reported a positive effect on PWA (28, 29). However, these studies were

Table 1. Demographic characteristics of the study population. [ISCO=International Standard Classification of Occupations]

Demographics	Participants (N=859)				Controls (N=2426)			
	Mean	Range	N	%	Mean	Range	N	%
Propensity score	0.109	0.002–0.615			0.093	0.002–0.623		
Gender								
Female			775	90			2182	90
Male			84	10			244	10
Age group (years)								
≤40			137	16			408	17
41–50			523	61			1549	64
≥51			199	23			469	19
ISCO grade								
1–2 (managers)			309	36			890	37
3 (technicians)			218	25			597	25
4 (clerks)			41	5			107	4
5 (service workers)			201	23			572	24
6–9 (manual workers)			90	10			260	11
Type of employer								
Municipality			587	68			1765	73
Hospital district			272	32			661	27
Residential region								
Southern Finland			493	57			1436	59
Central Finland			253	29			653	27
Northern Finland			113	13			337	14
Marital status								
Married/cohabiting			662	77			1889	78
Single/divorced/widowed			197	23			537	22
Educational level								
No high school			398	46			1086	45
High school			461	54			1340	55

Table 2. The prevalence and prevalence ratios of suboptimal perceived work ability (PWA) and mean score of PWA for the participants and controls at baseline and during a follow-up. [SD=standard deviation]

Time	Group	Prevalence of suboptimal PWA				Mean score of PWA		
		Prevalence (%)	Prevalence ratio	95% CI	P-value	Mean	SD	P-value
Baseline	Participants	22.1	1.03	0.89–1.19	0.713	8.15	1.29	0.776
	Controls	21.5	1.00			8.17	1.37	
Short-term [a]	Participants	33.2	1.23	1.10–1.39	<0.001	7.74	1.42	0.005
	Controls	26.9	1.00			7.91	1.59	
Long-term [b]	Participants	37.1	1.18	1.06–1.31	0.002	7.57	1.62	0.052
	Controls	31.4	1.00			7.71	1.84	

[a] The short-term follow-up of 1.7 (SD 1.02, range 0.003–4.55) years ended at the time of the 1st follow-up survey.
[b] The long-term follow-up of 5.8 (SD 1.14, range 3.01–9.16) years ended at the time of the 2nd follow-up survey.

limited to short periods of follow-up (up to 1 year) and small study samples (87–122 participants). The lack of effectiveness of VOMR found in our present study is in line with previous reports from this cohort, the finding being no reduction in the risk of work disability after VOMR (11, 13).

An inadequate selection of participants is a potential reason for an observed lack of effectiveness of VOMR. When rehabilitation is used as a measure of primary prevention, the identification of potential participants with an increased risk of future work disability is crucial. An individual need for rehabilitation (which is the main criterion for participant selection in conventional rehabilitation) is not applied in VOMR, which is based on the assumption that work-related strain is enough to lead to the deterioration of work ability (ie, every employee with a stressful job can be a potential rehabilitant even if no major health problems exist). Work-related strain is, however, a common finding among the working-age population, and it is not likely to be the main criterion for defining an individual's need for rehabilitation. Indeed, earlier findings suggest that participant selection to VOMR fails to recognize persons with a higher need for rehabilitation, and it seems to favor those with no risk of work disability (13).

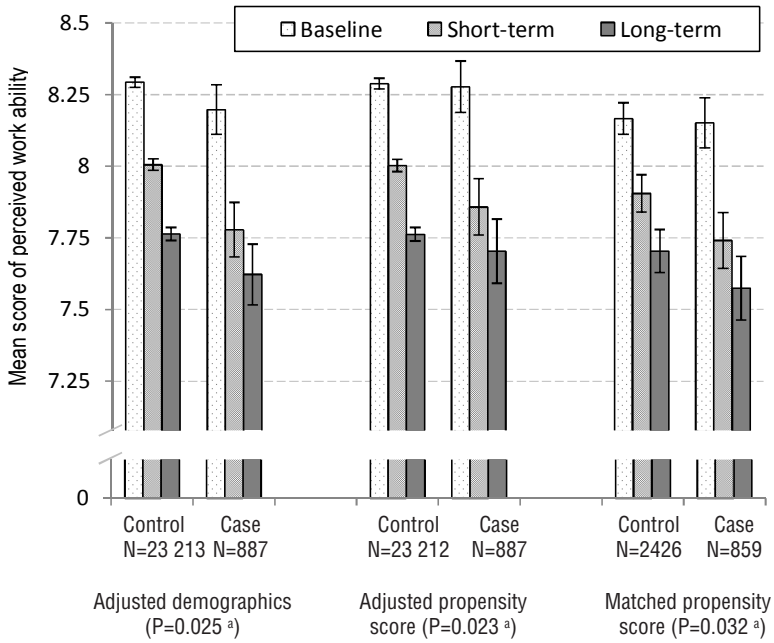

Figure 3. Mean scores and their 95% confidence intervals of perceived work ability at baseline and after the short - and long-term follow-up in the participants (case) and control group adjusted by demographics, adjusted by propensity scores and matched by propensity scores. The short-term follow-up of 1.7 [standard deviation (SD) 1.02, range 0.003–4.55) years ended at the time of the 1st follow-up survey. The long-term follow-up of 5.8 (SD 1.14, range 3.01–9.16) years ended at the time of the 2nd follow-up survey.

[a] Treatment year

The lack of effectiveness of VOMR may also be related to the individual-based nature of the program. Without a change at the participants' worksites, the individual-based prevention of strain may leave the causes of such strain unaffected (5, 6).

Participants do not work during the in-patient periods of rehabilitation, and employers are not financially compensated for hiring substitutes, potentially causing a financial burden on the employer and thus affecting negatively the relations between the supervisor and the employee. This situation could potentially have a negative impact on social relations in the workplace and increase work stress further, translating into a decline in work ability. However, the participants of our study came from the public sector, where such negative effects are not common. Moreover, it is employers and occupational health services that apply to the Social Insurance Institute of Finland for their employees to be placed in the VOMR program. Thus it is unlikely that the program causes a deteriorated PWA, a possibility that can partially explain our results. However, further studies are needed to investigate the contextual effects of the VOMR.

The main finding of the ineffectiveness of VOMR in relation to PWA was observed when a propensity-score-matched dataset was used in conjunction with adjustment of the dataset based on all the rehabilitants and non-rehabilitants. In this case, a decline of a similar magnitude was observed in the PWA mean score of the participants. However, in relation to the controls, this decline – though less steep than among the participants – was more obvious after adjustment than matching. We have found earlier that employees granted participation in VOMR displayed a lack of many important risk factors of work disability; the participants had, in fact, fewer behavioral health risks than the non-rehabilitants (13). This difference in the distribution of risk factors could be fully taken into account with the use of propensity-score matching, unraveling the poor performance of VOMR in terms of its effects on work disability, even better than more traditional comparisons based on adjustment.

We used the PWA score as an outcome for our study for two main reasons. First, this score has previously been found to correlate with work disability (21, 22), and, second, we have previously assessed and reported effects of the intervention on "hard" outcomes of work disability (11, 12). In the present study we were interested in assessing the effectiveness of VOMR from another point of view – participants' subjective work ability – as it is possible that workers do not always apply for (or get) sick leave even though they feel their work ability is declining.

The main strengths of the study are its (i) large sample size, (ii) use of national health registers with high coverage, (iii) repetition of surveys over time, (iv) matched control group, and (v) long follow-up. Although the distribution of the covariates used to derive the propensity score was the same for the cases and controls, propensity-score matching cannot remove bias due to unmeasured confounding when a strong selection bias exists. A common criticism on the usage of the pro-

pensity score is that if its determinants coincide with the determinants of the outcome measure, one could completely adjust to unity any true effect. However, this notion does not take into account that propensity score models include all variables that are needed to block the backdoor path between the exposure and the outcome. Therefore, it may also include predictors of the outcome. Still, the probability of residual bias due to unmeasured factors remains. However, when clinical indications and risks are similar, as in our study, strong selection bias and major confounding from unmeasured factors seem unlikely (20). Factors reflective of patient prognosis and physician decision-making behaviors are not available in observational datasets, although the likelihood of being treated depends on clinical judgment and referral selection. This situation is likely to result in an overestimate of the benefit due to residual confounding related to the selection of lower-risk patients for treatment; an underestimate would result from the selection of higher-risk participants (eg, for rehabilitation). However, propensity-score matching is likely to produce unbiased findings if the distribution of unmeasured prognostic factors is more likely to be similar when therapies with similar clinical indications and risk are considered. Under such conditions, randomized clinical trials and observational studies show the greatest similarities (30, 31). Because we focused on a low-risk population (13), strong selection bias did not occur in our study, and major confounding from unmeasured factors was unlikely. Another limitation of propensity-score matching is that it may lead to a loss of cases at the tails of the distribution of the propensity score, to the extent that they do not overlap. However, in this study, only 28 of the 887 rehabilitants, eligible for matching, were excluded due to a lack of a control subject. Moreover, analyses based on propensity scores may provide a more valid estimate of treatment effect than conventional observational studies that are based on multivariable adjustments (32). As the propensity-score matching is performed on a single calculated variable, it offers better control for bias from confounding and assures fewer drop-offs than conventional matching does. However, the possibility of confounding can never be ruled out in observational data.

The study population involved only employees in the public sector, and this limitation may have compromised the generalizability of our results. However, the occupational status of the participants varied widely, from managers to manual workers. Although previously found to be reliable (21, 23), the 11-point scale that we used may not have been sensitive enough to catch minor changes in the PWA of the participants.

Our results suggest that a vocationally oriented multidisciplinary rehabilitation program for employees without major medical conditions may be ineffective with respect to improving perceived work ability.

Acknowledgements

The Finnish Public Sector Study is supported by the Academy of Finland (projects 124271, 124322, and 132944) and the participating organizations.

References

1. Herbertsson TT, Orszag JM. The early retirement burden: assessing the costs of the continued prevalence of early retirement in OECD countries. Bonn: Institute for the Study of Labor; 2003.

2. Kuoppala J, Lamminpää A. Rehabilitation and work ability: a systematic literature review. J Rehabil Med. 2008 Nov;40(10):796–804.

3. Linton SJ. Early identification and intervention in the prevention of musculoskeletal pain. Am J Ind Med. 2002 May;41(5):433–42. http://dx.doi.org/10.1002/ajim.10052.

4. Loisel P, Buchbinder R, Hazard R, Keller R, Scheel I, van Tulder M, et al. Prevention of work disability due to musculoskeletal disorders: the challenge of implementing evidence. J Occup Rehabil. 2005 Dec;15(4):507–24. http://dx.doi.org/10.1007/s10926-005-8031-2.

5. Schaafsma F, Schonstein E, Ojajarvi A, Verbeek J. Physical conditioning programs for improving work outcomes among workers with back pain. Scand J Work Environ Health. 2011 Jan;37(1):1–5. http://dx.doi.org/10.5271/sjweh.3078.

6. Schaafsma F, Schonstein E, Whelan KM, Ulvestad E, Kenny DT, Verbeek JH. Physical conditioning programs for improving work outcomes in workers with back pain. Cochrane Database Syst Rev. 2010(1):CD001822.

7. Linton SJ, van Tulder MW. Preventive interventions for back and neck pain problems: what is the evidence? Spine (Phila Pa 1976). 2001 Apr 1;26(7):778–87.

8. Ebrahim S, Beswick A, Burke M, Davey Smith G. Multiple risk factor interventions for primary prevention of coronary heart disease. Cochrane Database Syst Rev. 2006(4):CD001561.

9. Marine A, Ruotsalainen J, Serra C, Verbeek J. Preventing occupational stress in healthcare workers. Cochrane Database Syst Rev. 2006(4):CD002892.

10. van Wyk BE, Pillay-Van Wyk V. Preventive staff-support interventions for health workers. Cochrane Database Syst Rev. 2010(3):CD003541.

11. Saltychev M, Laimi K, El-Metwally A, Oksanen T, Pentti J, Virtanen M, et al. Effectiveness of multidisciplinary primary prevention in decreasing the risk of work disability in low-risk population. Scand J Work Environ Health 2012;38(1):27–37. http://dx.doi.org/10.5271/sjweh.3169.

12. Saltychev M, Laimi K, Oksanen T, Pentti J, Virtanen M, Kivimaki M, et al. Effect of a multidisciplinary rehabilitation programme on perceived health among employees at increased risk of incapacity for work: a controlled study. Clin Rehabil. 2011 Dec 2. Epub ahead of print. http://dx.doi.

org/10.1177/0269215511425963.

13. Saltychev M, Laimi K, Oksanen T, Pentti J, Virtanen M, Kivimaki M, et al. Predictive factors of future participation in rehabilitation in the working population:The Finnish public sector study. J Rehabil Med. 2011 Apr;43(5):404–10. http://dx.doi.org/10.2340/16501977-0788.

14. OECD. Sickness, disability and work–keeping on track in the economic downturn. Stockholm: Organisation for Economic Co-operation and Development Directorate for Employment, Labour and Social Affairs; 2009.

15. SII. Statistical yearbook of the Social Insurance Institution, Finland, 2010. Helsinki: Social Insurance Institution of Finland; 2011. Available from: http://www.kela.fi/it/kelasto/kelasto.nsf/alias/Kunto_10_pdf/$File/Kunto_10.pdf?OpenElement.

16. Stuart EA. Matching methods for causal inference: A review and a look forward. Stat Sci. 2010 Feb 1;25(1):1–21. http://dx.doi.org/10.1214/09-STS313.

17. Huybrechts KF, Brookhart MA, Rothman KJ, Silliman RA, Gerhard T, Crystal S, et al. Comparison of different approaches to confounding adjustment in a study on the association of antipsychotic medication with mortality in older nursing home patients. Am J Epidemiol. 2011 Nov 1;174(9):1089–99. http://dx.doi.org/10.1093/aje/kwr213.

18. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70(1):41–55. http://dx.doi.org/10.1093/biomet/70.1.41.

19. Shah BR, Laupacis A, Hux JE, Austin PC. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. J Clin Epidemiol. 2005 Jun;58(6):550–9. http://dx.doi.org/10.1016/j.jclinepi.2004.10.016.

20. Stukel TA, Fisher ES, Wennberg DE, Alter DA, Gottlieb DJ, Vermeulen MJ. Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. JAMA. 2007 Jan 17;297(3):278–85. http://dx.doi.org/10.1001/jama.297.3.278.

21. Ahlstrom L, Grimby-Ekman A, Hagberg M, Dellve L. The work ability index and single-item question: associations with sick leave, symptoms, and health--a prospective study of women on long-term sick leave. Scand J Work Environ Health. 2010 Sep;36(5):404–12. http://dx.doi.org/10.5271/sjweh.2917.

22. van den Berg TI, Elders LA, de Zwart BC, Burdorf A. The effects of work-related and individual factors on the Work Ability Index: a systematic review. Occup Environ Med. 2009 Apr;66(4):211–20. http://dx.doi.org/10.1136/oem.2008.039883.

23. Martus P, Jakob O, Rose U, Seibt R, Freude G. A comparative analysis of the Work Ability Index. Occup Med (Lond). 2010 Oct;60(7):517–24. http://dx.doi.org/10.1093/occmed/kqq093.

24. Radkiewicz P, Widerszal-Bazyl M. Psychometric properties of Work Ability Index in the light of comparative survey study. International Congress Series. 2005(1280):304–9. http://dx.doi.org/10.1016/j.ics.2005.02.089.

25. Lipsitz SR, Kim K, Zhao L. Analysis of repeated categorical data using generalized estimating equations. Stat Med. 1994 Jun 15;13(11):1149–63. http://dx.doi.org/10.1002/sim.4780131106.

26. Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. Am J Epidemiol. 2005 Aug 1;162(3):199–200. http://dx.doi.org/10.1093/aje/kwi188.

27. Karazman R, Kloimüller I, Geissler H, Karazman-Morawetz I. Effects of ergonomic and health training on work interest, work ability and health in elderly public urban transport drivers. Int J of Ind Ergon. 2000;25(5):503–11. http://dx.doi.org/10.1016/S0169-8141(99)00037-2.

28. Leino P, Kivekäs J, Hänninen K. Effects of work-oriented fitness courses in lumberjacks with low back pain. J Occup Rehabil. 1994;4(2):67–76. http://dx.doi.org/10.1007/BF02110046.

29. Nevala-Puranen N. Effects of occupationally-oriented rehabilitation on farmers' work techniques, musculoskeletal symptoms, and work ability. J Occup Rehabil. 1996;6(3):191–200. http://dx.doi.org/10.1007/BF02110755.

30. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. N Engl J Med. 2000 Jun 22;342(25):1878–86. http://dx.doi.org/10.1056/NEJM200006223422506.

31. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med. 2000 Jun 22;342(25):1887–92. http://dx.doi.org/10.1056/NEJM200006223422507.

32. Austin PC. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. Stat Med. 2008 May 30;27(12):2037–49. http://dx.doi.org/10.1002/sism.3150.

Received for publication: 28 November 2011