

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Äikäs, Antti Hermanni; Absetz, Pilvikki; Hirvensalo, Mirja Hannele; Pronk, Nicolaas P.

Title: What Can You Achieve in Eight Years? : A Case Study on Participation, Effectiveness, and Overall Impact of a Comprehensive Workplace Health Promotion Program

Year: 2019

Version: Published version

Copyright: 2019 The Author(s)

Rights: CC BY 4.0

Rights url: <https://creativecommons.org/licenses/by/4.0/>

Please cite the original version:

Äikäs, A. H., Absetz, P., Hirvensalo, M. H., & Pronk, N. P. (2019). What Can You Achieve in Eight Years? : A Case Study on Participation, Effectiveness, and Overall Impact of a Comprehensive Workplace Health Promotion Program. *Journal of Occupational and Environmental Medicine*, 61(12), 964-977. <https://doi.org/10.1097/JOM.0000000000001699>

What Can You Achieve in 8 Years? A Case Study on Participation, Effectiveness, and Overall Impact of a Comprehensive Workplace Health Promotion Program

Antti Hermanni Äikäs, MSc, Pilvikki Absetz, PhD, Mirja Hannele Hirvensalo, PhD,
and Nicolaas P. Pronk, PhD

Objective: The aim of this study was to investigate participation and effectiveness of a multiyear comprehensive workplace health promotion (WHP) program. **Methods:** Participation and effectiveness data came from employer and vendor systems. Health data came from health risk assessments (HRA) and biometric screenings. Participation and effectiveness were analyzed using descriptive analyses, *t* tests, and Mann–Whitney *U* tests where appropriate. Overall impact was assessed using the PIPE Impact Metric. **Results:** Eighty-six percent of employees completed the HRA and 80% the biometrical screenings. Annual participation rate was 24%, and total reach was 58%. The portion of successful participants was 23% in 2010 to 2013 and 18% in 2014 to 2017. PIPE Impact scores were 18% for 2010 to 2013 and 14% for the 2014 to 2017 study periods. **Conclusion:** Despite modest annual participation rates, overall 8-year reach was considered reasonable. Conservatively, we consider the overall program impact to be moderate.

Keywords: effectiveness, implementation, participation, program design, program evaluation, workplace health promotion

A best practice worksite health promotion (WHP) program is a synthesis of an evidence-based design process, a cost-efficient implementation with ongoing evaluation, an engagement of an organization and its management and employees, and a successful process or support toward behavioral change, which in all leads to

better health and productivity outcomes at an individual and company level.^{1–6} Participation is a crucial element for a successful program to gain eligible health effects.^{6–8} Therefore, WHP programs should seek high participation levels. Descriptively, in a North American survey almost all US (89%) and Canadian (87%) employers (*n* = 335) state employee health and productivity as core components of their organizational health strategy, yet participation rates can be low in disease management or smoking cessation programs, often falling below 20%.⁹

The participation rates reported in systematic reviews vary a great deal. In a review of 24 studies, Bull et al found a wide range of participation rates between 8% and 97%, with a median of 61%.¹⁰ More recently, Robroek et al documented participation rates between 10% and 64% and a lower median of 33% in programs aimed at physical activity and/or nutrition.⁸ Even if single long-term programs have documented participation rates of over 70%,^{11,12} participation levels in WHP interventions tend to remain below 50%.^{8–10} The large variation among single studies may be explained by different employer sizes, mixed service contents, and varied duration and incentive usage, but also by the definition of participation.^{8,13} Some studies define the employees' intention to attend a program as participation, whereas others interpret an entry into a program as participation, and yet others define it only as long-term adherence.^{8,13}

It has been suggested that participation rates of completing a health risk assessment (HRA) should be investigated separately from attendance into the program's services.⁸ A mean participation rate of 57% (44.2% to 75%) for baseline HRA was calculated in a review of 37 studies.¹⁴ In North America, average participation in biometric screenings was reported to be 45% with incentives and 25% without incentives.⁹ One of the highest HRA rates have been achieved in the Johnson & Johnson multiyear program, where 76% of the eligible employees completed a health assessment both at baseline in 2002 and at follow-up in 2007.¹² This high participation rate and well-engaged population was largely a result of a significant financial incentive.¹² Research shows that incentives can push HRA participation from a 20% to 40% level to a level of 70% to 90%.^{15,16} The authors of the North American survey suggested that incentives are clearly effective in boosting a relatively simple action or task, such as completing an HRA, but financial incentives might not be enough to drive people to participate in programs that aim for long-lasting health behavior change.⁹

Initial participation rates lose their value, if dropout rates are high.¹⁷ Success in WHP efforts demands adherence that generates a sufficient number of participants with improved health.^{6,18,19} In an HRA review, Soler et al noted that the median retention rate for HRAs was 79%, meaning that every fifth employee dropped out before the follow-up assessment.¹⁵ In a review of 32 studies by Marshall et al, most of the physical activity interventions achieved retention rates from 51% to 63%.¹⁷ Long-term participation analyses are very rare, but one incentive-based WHP program revealed that participation rates lowered from 43% to 37% during 7 years, and 30% of the employees enrolled in the program continuously.²⁰

From the Faculty of Sport and Health Sciences, University of Jyväskylä, Mikkeli, Finland (Mr Äikäs); University of Eastern Finland, Institute of Public Health and Clinical Nutrition, Head of Research and Development, Provention Ltd., Finland (Dr Absetz); Faculty of Sport and Health Sciences, University of Jyväskylä, Mikkeli, Finland (Dr Hirvensalo); HealthPartners Institute and Chief Science Officer, HealthPartners, Minneapolis, Minnesota (Dr Pronk).

Funding: Metsämiesten säätiö, Finland, supported financially the manuscript writing process. Urheilupuistosäätiö, Finland, supported financially the data collection and analysis.

Conflicts of interest: AA is an employee of 4event Ltd, which designed and implemented the study's worksite health promotion program. PA has kept two single team workshops for 4event Ltd during 2013 to 2015; the participants (10 to 20 persons) were employee of Stora Enso Metsä. Other authors: None declared. Consultant Jouni Lahti, PhD, Adjunct Professor in the University of Helsinki: None declared.

Supplemental digital contents are available for this article. Direct URL citation appears in the printed text and is provided in the HTML and PDF versions of this article on the journal's Web site (www.joem.org).

Clinical significance: High participation rates (>80%) can be achieved in health risk assessments also without incentives. Eight-year worksite health promotion program (WHP) achieved reasonable total reach and moderate health impact, despite modest annual participation. Study's findings agree with earlier literature, suggesting that comprehensive WHP programs can promote lifestyle changes and health improvements.

Address correspondence to: Antti Hermanni Äikäs, MSc, Faculty of Sport and Health Sciences, University of Jyväskylä, Hauskankatu 1b, 50100 Mikkeli, Finland (anaikas@hotmail.com).

Copyright © 2019 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American College of Occupational and Environmental Medicine. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1097/JOM.0000000000001699

Another 7 year's investigation reported a total reach of 48%, but not dropout rates.²¹

A few simple reasons for the 20% to 40% dropout rate may be directly related to the workplace context. First, staff is not stationary, employees change jobs, business units and workplaces or they retire.^{6,9} Across all employee groups, turnover rates can range from 2% to 46%.²² According to a recent study, the annual average turnover rate was 22% in the United States and 20% in Canada.²³ The more the workforce switches, the more likely the decrease is in long-term adherence.⁴ Second, health screenings are usually offered to all employees, but "the need" for actual health promotion services reflects a smaller proportion of the staff.^{18,24} In a health risk reduction study, 13% of the participants were in a high risk group, 31% in a moderate risk group, and 56% in a low risk group.¹⁸ Considering this context, expecting to reach a high participation rate only from the employees who are at a moderate or high risk group may be overly optimistic.

Previously, some employee characteristics have been identified as being associated with participation, such as being familiar with health behavior and its benefits, concern for one's health, having a positive attitude toward the program, being ready to make changes, and giving high priority to the expected outcomes.^{24,25} Furthermore, women tend to participate more often than men, nonsmokers more often than smokers, white-collar workers more often than blue-collar workers, employees with day jobs more often than shift workers, and employees with secure jobs more often than part-time or temporary workers.^{8,11,25} It is not unexpected that employees with weaker intentions participate less.⁸ Unfortunately, employees with more severe health risks, such as an abnormal body mass index (BMI), are the least likely ones to participate.^{12,14,26,27} Programs that consist of several components have a moderate level of services, and programs where participation is possible during working hours are expected to have higher participation rates than those which have a minimal number of services, and where the events take place outside work.^{8,28}

To summarize the literature findings on participation, the participation rates can exceed 50% for HRAs, but tend to decrease or stay below 50% during intervention and long-term monitoring.^{8,14,20,21} Positive intentions and beliefs, as well as support received from peers, managements, and the environment, function as facilitators, whereas the factors related to time, health, and the job tend to function as barriers.^{8,24,25,27–29}

Although important, participation alone is not sufficient to produce population impact, but needs to be complemented with effective program strategies.²⁹ *Changes in behavior and improvement in health* on an individual level are necessary drivers for a program to be overall successful and effective.^{5,30} Several reviews have evaluated effectiveness of WHP programs from different types of perspectives. Soler (2010) evaluated programs including HRA, feedback, and a short-term (<1 year) intervention component (eg, health education, enhanced access to physical activity) and showed meaningful positive effect on a wide range of outcomes including decrease in tobacco use, alcohol use, seatbelt nonuse, dietary fat intake, blood pressure, cholesterol, health risk estimates, worker absenteeism, and health care service use, but not on fruit and vegetables intake, body composition, and cardiorespiratory fitness.¹⁴ Two systematic reviews evaluating physical activity interventions at workplace revealed modest effects at best on physical activity behavior,^{31,32} as well as on fitness, lipids, work attendance, and job stress.³²

In a systematic review, comprehensive WHP programs with duration ranging from few months up to 7 years, Osilla et al observed mixed results regarding the programs' impact on health-related behavior, substance use, physiological markers, and costs.³³ They also found that the evidence on absenteeism, incentive usage, and mental health was insufficient.³³ In a

systematic review of RCTs, where maximum duration was 2 years, the overall effect of the programs was perceived to be small but positive across work-related outcome measures such as perceived health, absence due to sickness, productivity at work, and workability.³⁴ Notably, effects were larger among younger populations and in interventions with weekly contacts, and smaller in studies which met the criteria of high-quality RCTs.³⁴

To summarize earlier literature, WHP programs *can* generate improvement in health behavior and in health parameters, reduce absenteeism, and thereby produce positive financial return for the employer,^{14,32,34,35} but clearly not all WHP efforts have been successful.^{30,31,33} Furthermore, very little is known about the long-term effectiveness of WHP programs, although a 7-year study reported savings from diminished health care costs and absenteeism,²⁰ a voluntary incentive-based program noted improvements in risk factors during 7 years of follow-up²¹ and a 6-year program yielded positive changes in health risks and ROI.¹²

Multiyear participation and effectiveness analyses are rarely reported in the literature.^{12,20,21} Therefore, the purpose of this study was to analyze participation levels of a comprehensive WHP program that lasted for 8 years (hereafter referred to as ENSO). The annual participants of the ENSO program were reported from the whole study population period and separate analyses were completed for three different health status groups (poor, moderate, good). In addition, cumulative 8-year participation analyses were done to represent the total reach of ENSO.

Earlier, ENSO was analyzed from a design and implementation perspective using the 4-S and Best Practice Dimensions developed by Pronk,^{2,36} and the content and implementation of the program's first 4 years were described in more detail.³⁷ In this study, we present a detailed description of the latter half of the ENSO program. The current study also pivots the evaluation perspective to observe "the generated returns of the program." To report ENSO's effectiveness, an analysis of successful participants and a comparison of characteristics of successful and unsuccessful participants were completed between the years 2010 to 2017, and separately for 2010 to 2013 and 2014 to 2017. In addition, the overall program impact was quantified using the PIPE Impact Metric Model based on target population penetration, implementation, participation, and effectiveness data.^{36,37}

METHODS

Study Design

This case study focuses on a multiyear implementation of a WHP in a single company, with a retrospective quasi-experimental study design without a control group at three different measurement points in 2010 to 2011, 2013 to 2014, and 2016 to 2017. In the area of health promotion, experimental and quasi-experimental designs are suitable for questions on the effectiveness outcome.³⁸ The participation and effectiveness analyses according to the PIPE Impact Metric Model by Pronk were carried out to reflect the generated health returns of the multiyear WHP executed in a real-world setting.³⁶ This model has been described in various levels of detail in previous discussions related to research translation,^{36,39} a systems approach and its dissemination and implementation in a real-world setting,^{36,39–41} and it has been used in evaluating diabetes prevention interventions^{42–44} and physical activity programs.^{36,45}

Intervention

Stora Enso Metsä, *the employer*, implemented a health promotion program (ENSO) in Finland as part of their global reThink transformation process starting in 2010 and lasting until 2017. The ENSO program can be classified as a comprehensive program, including all five elements that are based on the Healthy

People 2010 definition: health education, supportive environment, integration into organization's structure, linkage to related programs, and worksite screening.⁴⁶ The employer is a wood supply company, and its main business is to buy, harvest, and transport wood for Stora Enso mills throughout Finland. At the time, the employer had altogether over 100 company business units nationwide. Employees of Stora Enso Metsä were the participants of the program; subcontractors and family members were excluded. The ENSO program was a tailored version of a comprehensive WHP concept produced by the provider of the program, 4event Ltd.

As a theoretical foundation, ENSO deployed the transtheoretical model (TTM) and a mixture of different behavioral change techniques, such as self-monitoring and motivational interviewing.^{37,47,48} ENSO was executed in practice by the service provider's head coach and approximately 30 wellness coaches and professionals. The wellness professionals had at least a Master's degree and the coaches at minimum an undergraduate educational degree in either nutrition, physical education, health sciences, or coaching. The program management was shared mostly between human resources executives and the provider. A mutual pension company and occupational health care units were informed about the program, but they were not involved in the design or implementation processes until the very last year.

The main goal of the program was to improve the health and well-being of every employee. Altogether 27 different services took place from 2010 until 2013. The main aim of these services was to support low-effort, pleasant lifestyle changes and to create a positive health conception rather than a negative one. It is noteworthy that the emphasis of the services shifted in 2014. More services and tools were established to maintain and improve the workplace climate and stress management, as well as to strengthen mental resources from 2014 until 2017, whereas less targeted services related to nutrition, physical activity, and lifestyle were offered than during the first half. During these last 4 years, 49 different services were implemented. During its entire 8-year-long period, ENSO consisted of altogether 76 different services, three assessments of health risks, and seven annual WHP events, and it contained several communications materials such as the Vitality Book (2011) and six Service Books (2012 to 2017). A flowchart of the program during the years 2014 to 2017 is presented in Figure 1, and a more detailed information about the years 2009 to 2013 has been published earlier.³⁷

ENSO had six main components during its 8-year period. First, assessments of health risk with feedback were made available for all employees at three different time points, 2010 to 2011, 2013 to 2014, and 2016 to 2017. The assessment consisted of biometrical measurements and a HRA questionnaire. Biometrical screening was executed using a mobile Polar Body Age clinic, which was transferred into different locations to meet the employees.^{49,50}

As a second main component, based on the results of the HRA and biometrical screening, participants were offered targeted services of the 4event WHP concept. Support for lifestyle change was categorized into three different subgroups of health and fitness status: *poor*, *moderate*, and *good*. The participants were classified into the groups according to their Body Age measurement results. The result of the measurement is expressed as plus or minus years to be added to or reduced from the person's actual age. The result of the poor health status group was +6 years or higher, the moderate health status group between +1 and +5 years, and the good health status 0 year or less than person's own age. Furthermore, if an HRA participant was identified with a risk behavior or health risk (such as poor health status, physical inactivity, overweight, stress, functional insomnia, or fatigue), a coach invited the respondent to participate in suitable targeted service. Notably, the strongest life change support was offered to the population of poor health status with limited coaching places. Employees with moderate health status were encouraged to participate in group sessions with moderate

support, and employees with a good health status were offered group services with minimal lifestyle change support. The invitation outreach protocol included phone calls, text messages, group or individual emails depending on the size of the group, and reachability and data protection necessities of a service. Robust incentives were not used to promote participation into HRAs or other services. However, employees were free to attend some of the services during paid working hours.

The third main component of ENSO was that it had services that were offered to all employees. These included annual WHP events, open webinars, and hiking trips in Europe. In addition, the provider started local group coaching sessions in 16 locations in 2017.

As a fourth component, to ensure local awareness to the program within a widely dispersed organization, a playmaker network was established in 2011. Playmakers were nonmanagement employees volunteering to be trained by the provider to assist in the implementation and communications processes of ENSO. A playmaker was asked to promote well-being in their own local area. Every year they had their own list of tasks such as to enhance participation, promote workplace climate, or organize local low-effort sport events. A single playmaker was responsible for 15 to 30 employees, so to cover the entire staff, there were annually 23 to 28 playmakers in total.

With the help of the employer, the provider introduced several communication solutions, including home mailings, posters, a website, an intranet, service flyers, targeted phone calls, emails, Skype meetings, and electronic registration processes. This can be seen as the fifth main component of ENSO. It represented a communication strategy based on multichannel information to make the program as available and visible as possible to the employee audience.

Ongoing program management was ENSO's sixth component. This was a shared responsibility between the HR executives and representatives of the provider. Based on user experience, feedback, assessments of health risks, and the process evaluation reports, the provider planned the program's annual solutions and its continuum, and the HR executives made the final decisions. After reaching an agreement, the provider implemented the program. As stated in the design evaluation of ENSO, the process was lacking a plan for long-term sustainability,³⁷ and both parties were willing to optimize efficiency related to administrative work. As a result, to support the program management, the provider nominated a "head coach" to maintain the process, and deployed a customer relationship management system (CRM) to help with the management. In 2017, the service provider, the occupational health care provider, the mutual pension company, and the HR established a collaborative steering group which showed positive impact toward stronger cooperation at the end of the program.

Participants

The ENSO program and the assessments of health risk were made available for the whole staff of Stora Enso Metsä. Subcontractors were not involved. In the over 100 offices, most of the employees were executives, local forest officers, organization officials, and lumberjacks, whereas the staff in the working stations connected into Stora Enso's eight paper mills consisted of terminal workers. Statutory labor negotiations took place twice during the intervention. First, a new organization was established in 2013 to 2014 and as a result the total number of employees declined. After 2014, the local forest officers in Western Finland and the lumberjacks were no longer part of the organization. The second negotiation in 2017 did not have a major influence on the study population. Hence, the total number of employees decreased from 2010 to 2017 as follows: 651, 634, 630, 625, 530, 526, 523, and 523. Employees, who did not take part in HRAs or biometrical screenings (10% at baseline, 20% at the first follow-up, and 13% at the second follow-up), still retained access to the services during the whole program period of 2010 to 2017.

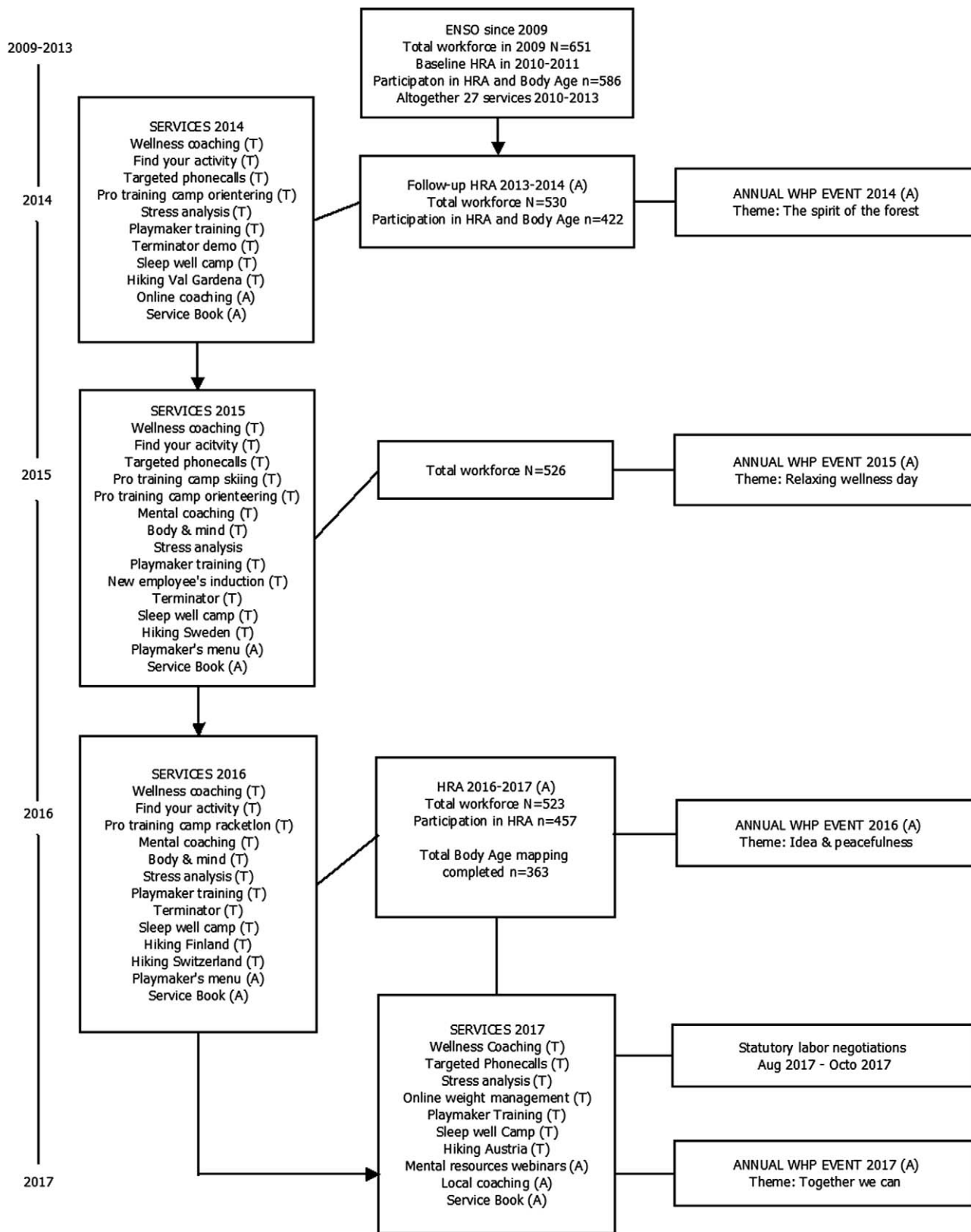


FIGURE 1. A flow chart of the program during 2014 to 2017. T = Targeted service for specific population based on location or content of the service; A = Service was available for all employees (see reference 37 for flow chart for the years 2009 to 2013).

Data Collection

There are a variety of ways to define *participation*, such as estimating intentions to use services or applications, enrolling into HRAs and into single services, completing coaching programs, and participating in self-care activities.^{6,24}

In this study, participation was defined as actual attendance at the activities that were implemented. The intention to participate or cases where an employee registered to a service but never took part were excluded. Information on participation was collected either from e-registration lists or from scanned lists of the participants' names.

The data about contacts (face-to-face vs e-contacts; group vs individual; total) were gathered by combining the data of actual attendance and the ordered number of contacts based on the provider's annual service books and orientation materials for the coaches. Each attendance to a service accumulated the contact data, which was used also in the effectiveness analysis.

Dropout rates from single services could not be used in this study, due to the fact that the majority of the services were executed face-to-face (f2f), and a separate list for those participants who left the service earlier than expected was not established by the employer or the service provider. In addition, registration to webinars and Skype meetings, as well as to the playmakers' local events, was more difficult to define, and this information was lost during the data collection. Therefore, the final participation data in our analysis were conducted along a priori defined parameters.

The final element of the PIPE Impact Metric Model is effectiveness. According to Pronk (2003), effectiveness refers to the rate of *successful participants*.³⁶ Effectiveness should be considered in the context of a program conducted in a real-world setting, and the criteria for success should be defined as part of the design phase: *what is planned to be achieved?*³⁶ Furthermore, criteria for success should be closely related to the design phase and to the program's ability to generate expected outcomes. Aziz et al, for example, associated the effectiveness of diabetes prevention programs with three main criteria: weight loss, diabetes risk reduction absolute, and relative.⁴² According to earlier findings of the ENSO program,^{37,51} the primary targets of the program were to support lifestyle changes, improve health metrics, and over time improve the health status of the people in poor or moderate health status groups. Based on these factors, participants were considered to be successful if they met two clear criteria. First being that the person answered YES to the question "Have you made a lifestyle change during the program?", and chose at least one from a list of nine lifestyle change options. The options were (1) I'm eating better quality fatty acids. (2) I eat more vegetables, salads, berries and fruits. (3) Nowadays, I enjoy a decent breakfast and lunch. (4) I have decreased the amount of high-caloric drinks and delicacies. (5) I have increased endurance training such as skiing, swimming, Nordic walking, cycling. (6) I have increased muscular fitness training such as gym, water aerobics, circuit training. (7) I have increased my everyday exercise. (8) I have decreased the use of alcohol. (9) I have cut down on smoking. Second, the person's biometric health parameters had to be improved based on the Body Age biometric screening.

The answers to the question about life change were collected with an HRA questionnaire⁵² at the same time with the Body Age biometrical screenings. In 2016 to 2017, employees who were not able to attend biometrical screenings filled the HRA questionnaire via a web-based survey tool. The questionnaire was a combination of questions from an annual national survey *Health Behaviour and Health among the Finnish Adult Population* (physical activity, habits),⁵³ Polar Body Age test protocol (testing safety), the stages of change transtheoretical model,⁴⁷ the employer's own questions (background information), and the provider's own questions (vitality, weight management, musculoskeletal disorders).⁵² The questionnaire also assessed absences due to musculoskeletal disorders, but not information concerning to health care costs or occupational health care usage.⁵²

The data on health improvement was based on the Body Age measurement. The Body Age method is a technology-aided testing system, primarily targeted for use in fitness and health-related environments to motivate individuals to be physically active and to improve their overall well-being^{49,50} and it was chosen by the provider to renew employer's physical health testing pattern.⁵¹ It contained five physiological factors and four performance factors.^{49,50} The five physiological factors were body mass index, body fat percentage, systolic blood pressure, diastolic blood pressure, and VO₂max. The four performance factors were number of crunches in 60 seconds, a leg endurance test, a bicep curl, and a sit-and-reach test. As a summation of each of these performance tests, the BodyAge system calculated the Body Age value, where poor health parameters increased the body age and good or excellent values decreased it. An example of two different Body Age values for same-aged males, is given in Figure 2. The method had been tested in a RTC study and more detailed descriptions of the measurement protocol and the Body Age calculation had been published earlier.^{49,54}

The Body Age value was calculated if a participant had at least four out of five physiological factors measured.⁴⁹ In this study, the health improvement was categorized as positive, if an employee succeeded to improve the overall evaluation of their Body Age. For example, if a participant had a result of +5 years in the baseline measurement and -2 years in the follow-up, the health improvement was counted as positive. If the Body Age stayed the same or got higher, it was considered as negative.

Most of the participation and effectiveness data were gathered as part of the service provision by the provider. To reduce potential bias in the data collection, the data resources, the participation lists, and the participation data collection were double-checked by an external WHP professional (PhD, Adjunct Professor at the University of Helsinki).

Data Analysis

Data on participation and effectiveness were analyzed from the whole program continuum 2010 to 2017. Descriptive statistics including frequencies, means, standard deviations, and percentages were used in reporting the characteristics of the study population and trends in participation. Participation percentages were calculated by dividing the number of actual attendants by the annual amount of employees. Additional participation analyses were conducted to represent the flow and accumulation of participation in the three different health status groups during the whole program. In these analyses, only those employees who took part in both the annual WHP event and the targeted services during the same year (hereafter referred to as BOTH group) were included.

Effectiveness analysis included those employees who completed two HRAs and biometrical screenings either between 2010 and 2017, 2010 and 2013 or between 2014 and 2017. To analyze the effectiveness rate among employees, a comparison of characteristics of successful and unsuccessful participants was conducted. The dichotomous variables were presented in percentages: life change (yes/no), health status (poor, moderate, good), gender, and personnel group, whereas continuous variables were presented in means and standard deviations. Statistical comparisons between age, Body Age change, and participation information were carried out by using the *t* test and the Mann-Whitney *U* test, as appropriate. A difference between the variables was considered statistically significant at a standard $P \leq 0.05$. All analyses were carried out using the IBM SPSS Statistics version 24.0 for Windows.

Finally, to complete the impact analysis, the total calculation of the PIPE Impact Metric score was calculated separately for the years 2010 to 2013 and 2014 to 2017 as follows: *penetration × implementation × participation × effectiveness*.³⁶ To estimate the *penetration* rate, both the provider and the employer gave their own independent evaluations about the amount of employees reached with invitations and communication materials, and the final penetration rate was an

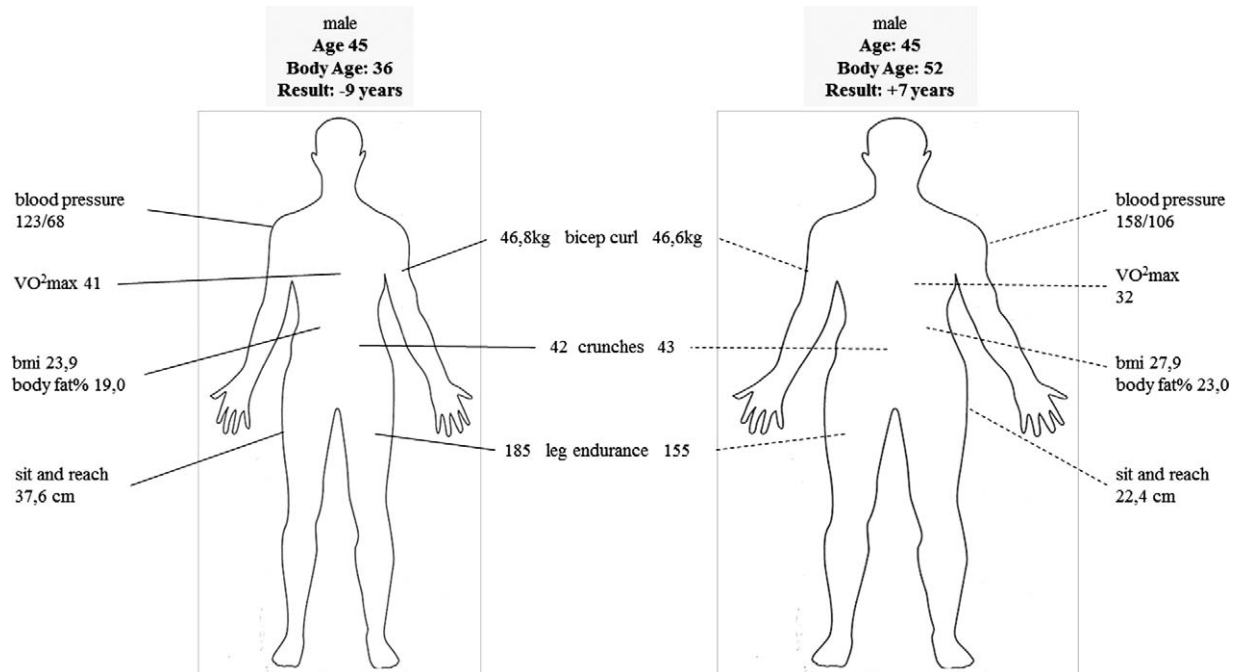


FIGURE 2. An example of two different Body Age values for the same age group and sex.

average of these two evaluations. The *implementation* coefficient represents the rate of implemented actions compared with the rate of planned actions.³⁶ The data for this factor were calculated similarly to Aikäs et al by comparing the provider’s annual budgets with the executed actions marked in the provider’s CRM database.³⁷ A coefficient for the *participation* was calculated by dividing the number of participants with the proportion of the target population that was reached with invitations (penetration).³⁶ In this study, the participation value for the PIPE Impact Metric’s analysis was a sum of BOTH participants. More specifically, the value for the years 2010 to 2013 represents the accumulative sum of participants who had participated in at least once into WHP event and targeted service either in 2011, 2012, or 2013. The participation rate for the years 2010 to 2017 and 2014 to 2017 was calculated the same way. In calculating the *effectiveness* coefficient, the number of individuals who met the two success criteria of the program were used as the numerator and the participation value was the denominator.³⁶

Ethical Issues

The employer and all the employees provided an informed consent for the program evaluation study. The protocol of the study with an informed consent form was included in the HRA remapping questionnaire both in 2013 to 2014 and in 2016 to 2017 and given to each employee. Participation in the study, as well as giving authorization to use earlier HRA and biometrical screening results and participation history, was voluntary. The data and the material related to the participation and effectiveness components of ENSO were retained carefully by the researcher. The study was conducted according to the ethical principles of the University of Jyväskylä and the research guidelines provided by the National Advisory Board on Research Ethics in Finland.⁵⁵

RESULTS

Participation in Assessments of Health Risk and Population Characteristics

The descriptive statistics of the representative sample of those who participated in the HRA and biometrical screenings

during the program are shown in Table 1. Participation in the baseline HRA questionnaire was 90%, 80% for the first follow-up, and 87% for the second follow-up (average 86%). The corresponding values for biometrical screenings were 90%, 80%, and 69% (average 80%). Both average participation rates were high. Most of the biometrical screening attendants were categorized into good health status group, 40% at the baseline, 47% at the first follow-up, and 51% at the second. The rates for moderate health status group at these assessment times were 29%, 29%, and 25%, respectively, and 31%, 24%, and 24% for poor health status, respectively. From a human resource perspective, all personnel groups were represented in the HRAs, although the number of lumberjacks decreased in 2013 and none of them was a part of the company after 2014 due to structural changes in the organization. Most of the HRA participants were male and local forest officers.

Participation

Of the total amount annual employees, 51% to 70% took part in annual WHP events in 2010 to 2017 and 16% to 70% in targeted services in 2011 to 2017 (see Table 2). The average percentage for attending the annual WHP events during the intervention was 63% and for targeted services 36%. The lowest attendance in targeted services was reported in 2011 (16%) and the highest in 2017 (70%). Most of the participants were male, 82% in the WHP events and 77% in the targeted services.

In this research, the *both participation* (BOTH) rate reflects the combined analysis of the WHP event and the targeted services. The annual amount of employees in the BOTH group varied between 14% and 42%, resulting as an average of 24%. Sex differences of the BOTH group followed the same pattern as in the earlier separated analysis, even though female employees tend to participate more often, if results (~74% males and ~26% females) are compared with the baseline sex distribution at the baseline (83%, 17%; see Table 1).

When studying the BOTH participation of the three different health status groups, the main finding indicated that every health status group was involved in the program each year. In the poor

TABLE 1. Participants in Assessments of Health Risks, Population Characteristics, and Descriptive Results at Three Different Time Points

Year	2010–2011	2013–2014	2016–2017
	N (%)	N (%)	N (%)
HRA participants	586 (90%)	422 (80%)	457* (87%)
HRA declined	65 (10%)	108 (20%)	66 (13%)
Sex			
Male	489 (83)	348 (82)	363 (79)
Female	97 (17)	74 (18)	94 (21)
Personnel group			
Executives	85 (14)	98 (23)	118 (26)
Local forest officers	275 (47)	207 (49)	224 (49)
Organization officials	110 (19)	59 (14)	62 (13)
Terminal workers	71 (12)	42 (10)	53 (12)
Lumberjacks	34 (6)	16 (4)	0
NA	11 (2)	0	0
	N = 586	N = 422	N = 363*
Participants in Biometrical Screenings	Mean (SD)	Mean (SD)	Mean (SD)
Age	43.8 (11.0)	45.0 (10.7)	44.8 (11.1)
BMI	26.9 (4.0)	26.6 (4.0)	26.8 (4.1)
Systolic BP, mm Hg	137 (16.6)	135 (15.5)	134 (16.3)
Diastolic BP, mm Hg	79 (10.1)	80 (9.7)	78 (9.6)
Body fat percentage	26.3 (6.7)	26.3 (6.5)	26.8 (6.4)
Bicep curl, kg	40.6 (10.3)	41.0 (10.5)	41.2 (10.8)
Sit and reach test, cm	32.2 (9.6)	36.4 (10.9)	36.6 (11.7)
Crunches in 60 s	36.8 (13.1)	32.0 (14.6)	33.0 (16.0)
Leg endurance, s	107 (53.1)	129 (60.7)	132 (64.2)
VO ₂ max, mL/kg/min	37.9 (8.3)	38.1 (8.2)	38.4 (8.3)
Fitness class 1–7	4.2 (1.3)	4.3 (1.3)	4.4 (1.3)
Health Status in HRA[†]	N (%)	N (%)	N (%)
Good	235 (40)	201 (47)	184 (51)
Moderate	168 (29)	121 (29)	91 (25)
Poor	183 (31)	100 (24)	88 (24)

*457 employees answered the HRA questionnaire and 363 of them completed also the biometrical screenings.

[†]Health status based on the Body Age (BA) calculations. If the BA was at most 0 years, a participant was categorized into the good health status group. The BA value between +1 and +5 years resulted moderate health status and the BA value at least +6 years was counted as poor health status.

health status group, 21 to 79 employees (28% to 56% of the total BOTH group) took part annually. Values for the moderate health status population were 12 to 74 employees (14% to 33%) and for the good health status group 25 to 81 employees (28% to 47%). An average participation rate for the 7 years was 32% for the poor, 26% for the moderate, and 38% for the good health status group. The number of participants in the poor health status group decreased to 20 to 35 people after the first 2 years, and rose up again in 2017.

BOTH rates were lower after the HRA in 2011, in 2013 to 2014, and also in 2016. The highest participation rates occurred in 2012 and 2017, when the program reached at least 38% of the employees.

In the personnel group analysis, most of the BOTH group's attendants (~55%) worked as local forest officers all over the country. Participation rates of executives (~13%) and organization officials (~22%) were proportional to total staff. Terminal workers (~6%) and lumberjacks (~3%), on the contrary, were not reached successfully.

The cumulative participation rate of the BOTH group within the whole study period is presented in Figure 3. Overall, the participation flow steadily increased across all health status groups. However, the number of people in the poor and moderate health

status groups increased slightly less than in the good health status group, and after 2013 those with the best health condition were the largest participant group. Taken together, 491 employees were ranked into the BOTH participation group at least once between the years 2010 and 2017. This result indicates that the ENSO program reached 58.1% (491/845) of those employees, who worked at least for 1 year in the organization. At the same time, the annual employee turnover rate varied between 3.6% and 17.0% with a mean of 7.8% per year (data available upon request). The highest accretions occurred during the same years as the highest participation jumps occurred: 2012, 2015, and 2017. A minority of the BOTH participants (n = 24) did not complete a HRA.

To summarize the participation results, the ENSO program achieved moderate gains in attendance every year after the launch of the program. Approximately one in four (24%) employees took part in the program annually, and with the help of long-term implementation, the program managed to reach 58% of the total workforce at least once during the 8 years. Most of the participants were males which represented the study population. There were no major attendance differences among the three different health status groups. The only personnel groups which were not deeply involved in the program were terminal workers and lumberjacks.

TABLE 2. Participation Rates in WHP Events, Targeted Services, and Separated BOTH Participation Analysis During 2010 to 2017

Year	2010	2011	2012	2013	2014	2015	2016	2017	Mean%
Total personnel N	651	634	630	625	530	526	523	523	
	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	63%
Partic. in WHP event	431 (66%)	423 (67%)	428 (68%)	437 (70%)	289 (55%)	367 (70%)	269 (51%)	312 (60%)	
Sex									
Male	361 (84)	351 (83)	360 (84)	363 (83)	225 (78)	307 (84)	223 (83)	243 (78)	82
Female	70 (16)	72 (17)	68 (16)	74 (17)	64 (22)	60 (16)	46 (17)	69 (22)	18
Partic. in targeted services		104 (16)	268 (43)	130 (21)	132 (25)	223 (42)	189 (36)	365 (70)	36%
Sex									
Male		79 (76)	216 (81)	96 (74)	91 (69)	173 (78)	150 (79)	294 (80)	77
Female		25 (24)	52 (19)	34 (26)	41 (31)	50 (22)	39 (21)	71 (20)	23
BOTH participation (WHP event and TS)		88 (14%)	240 (38%)	95 (15%)	90 (17%)	148 (28%)	76 (15%)	222 (42%)	24%
Sex									
Male		68 (77)	195 (81)	67 (71)	58 (64)	114 (77)	54 (71)	175 (79)	74
Female		20 (23)	45 (19)	28 (29)	32 (36)	34 (23)	22 (29)	47 (21)	26
Health status*									
Good		25 (28)	81 (34)	39 (41)	42 (47)	64 (43)	28 (37)	86 (39)	38
Moderate		12 (14)	74 (31)	21 (22)	20 (22)	44 (30)	25 (33)	64 (29)	26
Poor		49 (56)	79 (33)	27 (28)	26 (29)	35 (24)	21 (28)	65 (29)	32
NA		2 (2)	6 (3)	8 (9)	2 (2)	5 (3)	2 (2)	7 (3)	4
Personnel group									
Executives		9 (10)	33 (14)	9 (10)	19 (21)	24 (16)	5 (7)	28 (13)	13
Local forest officers		57 (65)	122 (51)	47 (49)	52 (58)	61 (41)	48 (63)	126 (57)	55
Organization officials		19 (22)	48 (20)	20 (21)	19 (21)	26 (18)	22 (29)	49 (22)	22
Terminal workers		1 (1)	20 (8)	1 (1)	0 (0)	37 (25)	1 (1)	16 (7)	6
Lumberjacks		2 (2)	13 (5)	15 (16)	NA	NA	NA	NA	3
NA		0 (0)	4 (2)	3 (3)	0 (0)	0 (0)	0 (0)	3 (1)	1

*The health categorization was based on the participant's first HRA. For example, if an employee participated in the services in 2011, but had her/his first HRA in 2013 and the second in 2017, this analysis used the health status from 2013 assessment.

Effectiveness

Table 3 provides a comparison of successful and unsuccessful participants from 2010 to 2017, 2010 to 2013, and 2014 to 2017. To be included in the effectiveness analysis, a participant had to have two HRAs and biometrical screenings completed from the 2010 to 2017, 2010 to 2013, or 2014 to 2017 time periods. The inclusion gathered 253 people into the whole continuum, 359 people into the first time period, and 255 into the last. Notably, 215 participants completed all three HRAs and biometrical screenings, meaning that a large part of the participants were included in all periods' analysis (more detailed data is presented in Appendix 1, <http://links.lww.com/JOM/A599>). Of included participants, 134 persons in 2010 to 2017, 143 persons in 2010 to 2013, and 96 persons in 2014 to 2017 met the criteria of success (lifestyle change and improvement in health) during the intervention. The successful rates for participants were 52% (134/253) and 40% (143/359) and 38% (96/255) for the same three periods, respectively. Altogether 67% of the HRA participants (281/422) reported that they had made a lifestyle change during the first half of the intervention. During the second half, the proportion was 55% (198/363) and 87% (220/253) for those who completed both the baseline HRA and the last follow-up HRA.

There were no demographic (age, sex) differences between the successful and the unsuccessful groups. When comparing health status (good, moderate, poor) in the successful group, the differences of successful participants were quite similar during all periods, 34%, 36%, and 25% of the successful participants came from the poor health status group. Corresponding results for the moderate health status group were 24%, 29%, and 37.5%. For the good health group, the rates were 42%, 35%, and 37.5%,

respectively. Noticeably, most of the unsuccessful participants came from the good health status group: 51% in 2010 to 2017 and 2010 to 2013 and 59% in 2014 to 2017.

In the personnel group analysis, local forest officers composed the majority group behind successfulness, followed by organization officials and executives. Terminal workers and lumberjacks showed the lowest success rates in all investigations periods.

There were notable differences in participation and contacts between the groups of successful and unsuccessful participants. Participation in annual WHP events did contribute to effectiveness only during the whole program period ($P = 0.040$), but not during the halves of the program in our analysis ($P = 0.108$, $P = 0.147$). At the early years of the intervention, participation in targeted services tended to be higher in the successful group ($P = 0.030$), but the difference was not significant during 2010 to 2017 or 2014 to 2017. The BOTH participation was slightly higher in the successful group between 2010 to 2017 and 2014 to 2017 than in the unsuccessful group ($P = 0.020$, $P = 0.034$). When comparing the contacts between the groups, primary findings were that the successful group received more face-to-face meetings and group contacts in all three study periods ($P < 0.010$). Correspondingly, the total amount of contacts was somewhat higher in the successful group than in the unsuccessful group ($P = 0.034$, $P = 0.014$, $P = 0.044$). Interestingly, the groups did not differ in e-contacts or individual contacts.

Pipe Impact Metric Score

PIPE Impact Metric score was calculated as a product of both design and execution elements of the WHP program. The calculation formula along with the coefficients of each of the four elements and the resulting PIPE Impact Metric score for the Enso program for

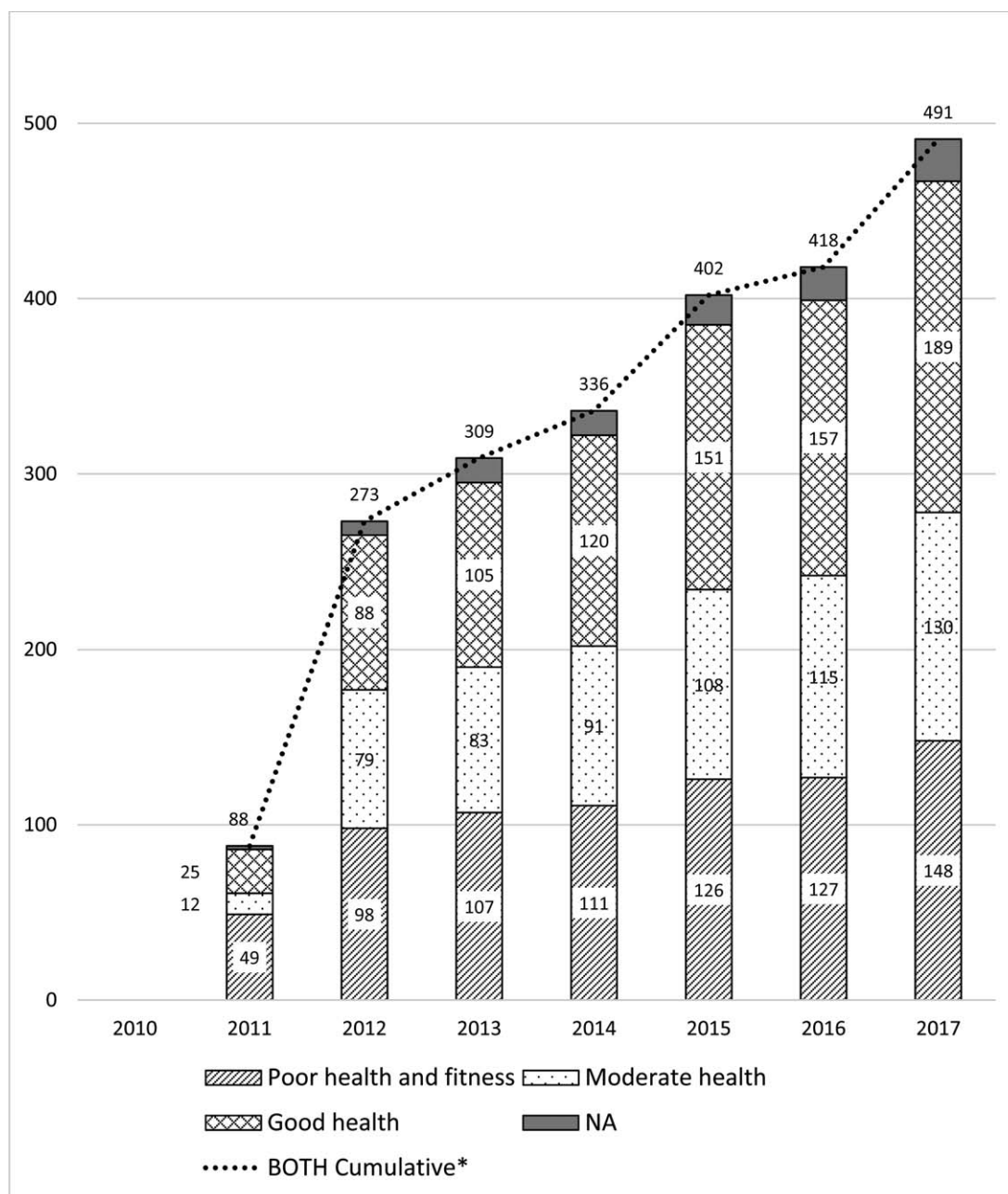


FIGURE 3. Cumulative BOTH Participation in the three health status groups from 2010 to 2017. *The BOTH participation denoted employees who took part both into WHP event and targeted service(s) during the same year.

years 2010 to 2013, 2014 to 2017, and 2010 to 2017 are presented in Figure 4.

The first half of the ENSO program reached the total coefficient of 0.182 (18.2%), which derived from high penetration (0.992) and implementation (0.808) rates accompanied with moderate participation (0.490) and effectiveness (0.463) rates.

In the second half of the program, the coefficient was 0.143 (14.3%). Compared with 2010 to 2013, penetration (0.952), implementation (0.784), and effectiveness rates (0.279) decreased, whereas participation rate was little higher (0.687) than in the earlier period. More detailed calculations of penetration and

implementation rates for 2010 to 2013 and 2014 to 2017 are presented in appendixes 2 (<http://links.lww.com/JOM/A600>), 3 (<http://links.lww.com/JOM/A601>), 4 (<http://links.lww.com/JOM/A602>), and 5 (<http://links.lww.com/JOM/A603>). The total coefficient for the whole program continuum was derived analysis of the halves and resulted level of 0.164 (16.4%). However, this numerical value contained overlapping from the years' 2010 to 2013 and 2014 to 2017 penetration, implementation, and participation components.

As an overview for the effectiveness analysis, the first half of the ENSO program was more effective than the latter half. Approximately one-fifth (23% and 18%) of the whole staff achieved the

TABLE 3. A Comparison of Successful and Unsuccessful Participants Between 2010 and 2017, 2010 and 2013, and 2014 and 2017

	Years 2010–2017			Years 2010–2013			Years 2014–2017		
	Successful	Unsuccessful	P	Successful	Unsuccessful	P	Successful	Unsuccessful	P
N	134	119		143	216		96	159	
Mean age (SD)	47.0 ± 9.7	48.4 ± 10.0	0.259*	44.3 ± 10.2	47.0 ± 10.2	0.216*	46.8 ± 10.5	46.7 ± 10.3	0.984*
Mean Body Age change (SD)	-4.5 ± 2.9	+1.7 ± 3.5	0.000*	-4.2 ± 2.5	+1.3 ± 3.3	0.000*	-3.7 ± 2.5	+1.6 ± 3.5	0.000*
Sex, N (%)									
Male	106 (79%)	107 (90%)		110 (77%)	191 (88%)		78 (81%)	135 (85%)	
Female	28 (21%)	12 (10%)		33 (23%)	25 (12%)		18 (19%)	23 (15%)	
Life change, N (%)									
Yes	134 (100)	86 (72)		143 (100)	138 (64)		96 (100)	102 (64)	
No		33 (28)			78 (36)			57 (36)	
Health status									
Good	56 (42)	61 (51)		50 (35)	110 (51)		36 (37.5)	93 (59)	
Moderate	33 (24)	34 (29)		42 (29)	61 (28)		36 (37.5)	40 (25)	
Poor	45 (34)	28 (20)		51 (36)	45 (21)		24 (25)	26 (16)	
Personnel group, N (%)									
Executives	29 (21)	21 (18)		21 (15)	38 (18)		20 (21)	34 (21)	
Local forest officers	85 (64)	69 (58)		72 (50)	112 (52)		58 (60)	81 (51)	
Organization officials	11 (8)	7 (6)		35 (25)	33 (15)		10 (10)	23 (15)	
Terminal workers	9 (7)	22 (19)		11 (8)	23 (11)		8 (9)	21 (13)	
Lumberjacks	—	—		4 (3)	10 (5)		—	—	
Participation	Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)	
In annual WHP events	6.2 ± 1.7	5.5 ± 2.2	0.040 [†]	3.2 ± 1.0	3.1 ± 1.1	0.108 [†]	3.0 ± 1.1	2.8 ± 1.3	0.147 [†]
Targeted services	2.8 ± 1.8	2.7 ± 1.7	0.778 [†]	1.1 ± 1.0	0.8 ± 0.9	0.030 [†]	2.0 ± 1.1	1.8 ± 1.2	0.408 [†]
BOTH	2.3 ± 1.6	1.8 ± 1.4	0.020 [†]	0.9 ± 0.9	0.7 ± 0.8	0.083 [†]	1.5 ± 1.0	1.3 ± 1.1	0.034 [†]
Contacts									
f2f	14.3 ± 9.4	10.9 ± 6.5	0.004 [†]	6.1 ± 3.8	5.2 ± 3.7	0.007 [†]	8.4 ± 5.6	7.1 ± 6.4	0.009 [†]
e-contacts	5.1 ± 5.5	4.6 ± 4.4	0.888 [†]	2.4 ± 3.4	1.9 ± 3.1	0.090 [†]	3.0 ± 3.7	3.0 ± 3.5	0.906 [†]
Group	14.6 ± 9.3	11.4 ± 6.5	0.006 [†]	6.0 ± 3.6	5.1 ± 3.5	0.006 [†]	9.0 ± 5.9	7.5 ± 6.4	0.006 [†]
Individual	4.9 ± 5.1	4.1 ± 3.9	0.676 [†]	2.5 ± 3.4	1.9 ± 3.0	0.072 [†]	2.6 ± 3.6	2.6 ± 3.3	0.914 [†]
Total contacts	19.2 ± 12.1	15.4 ± 8.0	0.034 [†]	8.6 ± 5.9	7.0 ± 5.3	0.014 [†]	11.2 ± 7.3	10.0 ± 7.9	0.044 [†]

*P values are based on the independent samples of *t* test.

[†]P values are based on the independent samples of Mann–Whitney *U* test.

success criteria (143/630 in 2010 to 2013 and 96/523 in 2014 to 2017).

DISCUSSION

This study investigated the degrees of participation and effectiveness in a multiyear comprehensive WHP program. In addition, this research completes the estimate of the overall impact of an 8-year-long design and implementation process of a comprehensive WHP program.³⁷ The analysis revealed that among a male-dominated workforce in dispersed organization, a comprehensive program without robust incentives achieved reasonable, long-term participation rates and moderate success rates. Conservatively estimated, the program’s degree of penetration into the workforce, implementation levels, and participation and effectiveness rates generated a moderate overall impact.

The participation findings of the present study are in line with the earlier literature, suggesting that attendance levels tend to stay below 50%.^{8,9,21} However, our study might underestimate annual engagement into the program due to the following reasons. First, to achieve participation status, an employee had to attend *both* an annual WHP event and the targeted services. Second, the more ambiguous data of webinars, Skype meetings, and playmakers’ local events were excluded in this study. In this regard, the participation result might be understated.

The HRA questionnaires and biometrical screenings reached a high number of participants in each three time points, the average being 86% and 80% without incentives. This contradicts the findings of earlier

studies, suggesting that financial stimulus is needed to gain a high recruitment level for health assessments.^{9,11,12} In this program, the HRA questionnaire and the Body Age screening were made convenient to complete,^{50,52} those were offered close to the participants’ workplaces and meetings, at multiple times throughout the day so as to fit into work schedules to complete the mapping during paid working hours.

We did not find any major gaps or unexpected drops in adherence in our investigation, even though our earlier study revealed challenges for sustainability in the design phase.³⁷ BOTH participation were lower in 2011, in 2013 to 2014, and in 2016. On the contrary, the highest participation rates occurred in 2012 and 2017. Following observations explained variation in BOTH participation. When targeted services were established in 2011, most of the support was offered to the poor health status group accompanied by a narrow selection of other services. In 2013 to 2014, statutory labor negotiations took place in Stora Enso Metsä, and it most likely hindered participation into WHP efforts. A computational reason behind the lower BOTH rates in 2014 and 2016 was the fact that then the number of employees taking part in the annual WHP events was lower than the program’s average. Respectively, during the second year (2012) of the program, several different coaching groups were established to engage more employees into the program and in 2017, a set of nationwide local coaching sessions was organized in 16 strategic locations to reach the maximum number of employees.

The majority of employees irrespective of their health status (poor, moderate, good) attended every year, suggesting that the

	PENETRATION		IMPLEMENTATION		PARTICIPATION		EFFECTIVENESS	
Pipe Impact Coefficient	$\frac{n \text{ (outreach)}}{N \text{ (target population)}}$	X	$\frac{\Sigma \text{ (executed services)}}{\Sigma \text{ (planned services)}}$	X	$\frac{n \text{ (participated)}}{n \text{ (outreach)}}$	X	$\frac{n \text{ (successful participants)}}{n \text{ (participated)}}$	= Π
2010-2013	$\frac{n \text{ (630)}}{N \text{ (635)}}$	X	$\frac{\Sigma \text{ (59)}}{\Sigma \text{ (73)}}$	X	$\frac{n \text{ (309)}}{n \text{ (630)}}$	X	$\frac{n \text{ (143)}}{n \text{ (309)}}$	= 0.182
Subscores	0.992		0.808		0.490		0.463	
2014-2017	$\frac{n \text{ (501)}}{N \text{ (526)}}$	X	$\frac{\Sigma \text{ (69)}}{\Sigma \text{ (88)}}$	X	$\frac{n \text{ (344)}}{n \text{ (501)}}$	X	$\frac{n \text{ (96)}}{n \text{ (344)}}$	= 0.143
Subscores	0.952		0.784		0.687		0.279	
2010-2017	$\frac{(0.992 + 0.952)}{2}$	X	$\frac{\Sigma \text{ (128)}}{\Sigma \text{ (161)}}$	X	$\frac{n \text{ (491)}}{n \text{ (630)}}$	X	$\frac{n \text{ (134)}}{n \text{ (491)}}$	= 0.164
Subscores	0.972		0.795		0.779		0.273	

FIGURE 4. PIPE Impact Metric scores in 2010 to 2013, 2014 to 2017, and 2010 to 2017.

concept, including the HRA and the provider’s services, found its targets. After the year 2013, most of the participants belonged to the good health status group, which corresponds with the same group’s prevalence in the HRAs. One reason why people in poor and moderate health status groups did not participate as often as in the early years of the intervention might be the fact that the emphasis of the program changed in 2014 to support workplace climate and mental resources, and there were less targeted services to support lifestyle change among poor and moderate health status group than during the first half.

In the relevant literature, it has been suggested that a moderate level of services and opportunities to participate during working hours are associated with higher participation rates,²⁹ and that multicomponent interventions are expected to achieve higher attendance rates than interventions with less components.⁸ Even though we found (only) a modest annual attendance, the cumulative participation analysis showed that new participants were recruited into the program every year. This supports the idea that addressing multiple “entry channels” was the right decision at the design phase, a conclusions from several other studies as well.^{2,6,26,29}

The groups that could not be reached properly were the lumberjacks and the employees in the terminal units. Lumberjacks worked mainly by themselves in the forest, where as a three-shift work was dominant for terminal workers. The work environment might have influenced both groups’ participation intentions. When comparing participation and sex distribution of the total staff, female employees attended the services slightly more often than males. Overall, our results corroborate previous results, indicating that relatively speaking, female workers, white-collar workers, and employees without shift work tend to engage in the programs more often.^{8,11,25,26}

What remains unclear, however, is how much the range of over 100 different business units affected the participation. The issue was not investigated in this study, but it seems evident that the participation cannot rise high if the distance to the service is a major barrier or if attending otherwise requires high effort. This notion is supported by the fact that the highest attendance into targeted services was seen in 2017, when local easy-to-go coaching sessions were established.

The primary targets of the program were to support participants in making lifestyle changes and to improve their health metrics.^{37,51} Approximately one in five of the total staff met the double criteria for a successful change, and more than half of the HRA participants reported that they had made a lifestyle either during the first (67%), during the last half (55%), and (87%) during the whole continuum. Success rates or proportions are seldom reported in WHP literature, but some earlier investigations offer insights. A long-term incentive-based program noted the following positive changes in a 3700 employees cohort: -15.5% in physical inactivity, -11.9% in poor nutrition, -3.5% in smoking, and +7.7% in safety belt usage and +7.5% in the proportion of low-risk employees. However, results did not report on positive changes in the proportion of employees reducing their high cholesterol, hypertension, or BMI.²¹ In a single year’s follow-up study, 47% of hypertension intervention’s participants self-reported an increase in physical activity and 92% of participants reported they tried to improve their diet.⁵⁶ As a result, 19% of participants achieved goals for vigorous physical activity and positive changes in weight and blood pressure were observed among the experimental group as compared with the control group.⁵⁶ Another multiyear study reported that among individuals who completed a weight control program, over 50% reduced their BMI, although the permanent effect faded on fourth year.⁵⁷ Another intervention targeting obese employees noted that 18% of participants managed to lower their Framingham-based CHD risk scores.⁵⁸ This study’s findings corroborated earlier literature,^{21,56-58} suggesting that self-reported behavior changes might reach higher proportions, but long-term improvements in health parameters are harder to achieve. In this study, a reason behind the result might be that in the data collection, a strict level for a required lifestyle change was not set, only the nine focusable lifestyle change options were given, and therefore some lifestyle changes might not have been sufficient enough to contribute to the biometrical health benefits, or they might have been contradicted by other, unhealthy changes. Another point is that the bar for success was set very high in this investigation. An employee needed to self-report success, make improvements, and maintain nine different health parameters for 3 to 4 years. For example, a

person who self-reported a life change, increased exercise, and improved VO₂max, but failed to improve the total Body Age index, was not counted as successful. However, we decided to set the bar high in the evaluation because the basic idea of the PIPE Impact Metric is to compare the program outcomes with its aims.

Furthermore, an additional reason for why the health improvement was not easy to overcome is that a majority of the unsuccessful participants (51% to 59%) belonged to the good health status group, and the main focus and the program's behavior change support was not targeted at this group.³⁷ It is also evident that if a participant has a high health parameter level already from the start, it is not easy to gain improvements in the follow-ups.

We found significant differences between the successful and the unsuccessful groups. The primary finding was that the successful groups received more face-to-face, group, and total contacts in 2010 to 2017, 2010 to 2013, and 2014 to 2017. On the contrary, the amount of e-contacts did not differ between groups. Interestingly, neither did the amount of individual contacts.

Based on the current findings, we found two important notions. First, face-to-face contacts are most likely needed to enhance long-term health improvements in WHP interventions, and even if e-contacts are easy to execute in the digitalized world, they alone might not be enough to make the program effective. Second, contrary to our expectations, the amount of individual support did not have a strong impact on the effectiveness of the program. Similar findings have been reported by Elliot et al and MacKinnon et al, who found team-based health promotion classes having more positive influence on physical activity than individual counselling.^{59,60} During this study's first half, the individual contacts entailed mainly coaching for physical health, and they were targeted to the poor health status population, of whom not all managed to reach the bar of success. During the latter half of the program, the emphasis of the targeted services changed and the individual contacts focused more on stress management and mental resources, and consequently the individual contacts did not help to achieve an improvement in this study's physical health parameters.

From a quality improvement perspective, the impact of individual contacts should have been maximized or the group contacts should have been given more emphasis. Like stated before, the role of the occupational health care units was minimal in the design and implementation phases.³⁷ Most likely, a closer cooperation and surveillance by health professionals especially in the poor health status group might have helped them to achieve and maintain better results after individual contacts.

We used the PIPE Impact Metric model as our evaluation tool. Basically, the PIPE (penetration, implementation, participation, and effectiveness) is an analysis method that shows the impact of the measured factors on the health of the population, and offers a scientific framework and feedback loops for quality improvement.³⁶ The impact quantified in our research reached 18% of the participants during the first half of the study, and 14% during second half of the study. There are no earlier PIPE calculations in the area of WHP to compare with, yet an example of the 10,000 steps' walking program for health plan members with diabetes mellitus resulted in a total impact of 6%^{36,45} and a peer support diabetes program achieved 1.7% level.⁴³ In a review of diabetes interventions, the PIPE Impact Metric elements penetration, implementation, and participation scores were categorized according to a three-level scale with following limit values: ≤33% = low, 34% to 66% moderate, and ≥67% high.⁴² The corresponding values for effectiveness scores were ≤25% = low, 26% to 40% moderate, and ≥40% high.⁴² If the same scale had been used in the present study, the related categories would have been *high* for penetration, *high* for implementation, *moderate* (2010 to 2013) and *high* (2014 to 2017) for participation, and *high* (2010 to 2013) and *moderate* (2014 to 2017) for effectiveness. However, the scale used by Aziz et al for

diabetes prevention programs may need specifications when applied to an overall WHP setting.

What do the percentages, 18% and 14%, tell about the level of impact in this study? Considering our result in context of the four different PIPE components, we estimate benchmark values from previous health promotion literature as follows: most interventions achieve higher than 90% penetration rate.^{37,42,45} The implementation rate typically varies between 60% and 90%,⁶¹ the participation rate tends to be below 50%,^{8,9} and the majority of the population will not achieve the program goals.^{21,31–33,57} Choosing liberal estimates based on the earlier literature of 95%, 85%, 45%, and 35% for the four elements, respectively, and transferring them into the PIPE Impact Metric calculation sheet, we would get a result of 13% (0.95 × 0.85 × 0.45 × 0.35). In this context, coefficients from our study may be considered as relatively high. However, as a final synthesis of earlier PIPE Impact Metric evaluations^{37,42–45} and relevant WHP literature,^{8,9,14,31–33,57} we consider the overall impact of the program to be moderate.

Furthermore, even though the percentages, 14 and 18, might appear to be small numerical values, it should be noted that in the absence of a health promotion program, the tendency is that the employees increase their risk factors over time (this would be represented by a negative PIPE Impact score).⁶² Thus, maintaining health (not increasing risks) is a good result, and it would be represented by a "0" PIPE Impact score. In this context, a PIPE Impact score of 14% or 18% over the course of 8 years may be considered a relatively strong impact on health.^{18,62}

In our calculations, the participation and effectiveness coefficients were clearly the weakest points of the program. The annual BOTH participation rates varied between 14% and 42% and over the course of its 8 years, the program engaged over 58% of all the employees. These numbers seem to be reasonable, but provide insight into opportunities for improvement in the recruiting and communication components of its design. Another challenge was the number of successful participants. As stated before, the effectiveness bar was set high in this investigation. If, for example, we had limited the effectiveness definition solely to the number of employees who self-reported that they had made a life change, the numerators for the effectiveness analysis 2010 to 2013 and 2010 to 2014 would have been 281 and 198 instead of the current 143 and 96, thereby doubling the number of successful participants and effectiveness of the program.

Notable strengths of this study include its long duration and a clear focus to investigate the generated health returns of the multi-year program in a real-world context. Balanced against these strengths are several limitations. First, to the best of our knowledge, this was the very first multiyear investigation on participation, effectiveness, and total impact of WHP program where the PIPE Impact Metric framework was utilized. This limits benchmarking and generalization of results. Second, our case study did not have a control or comparison group, and it is possible that during the 8 years there might have been confounders that our data collection and analysis could not reveal. For example, a single local team could have established their own wellness coaching projects via recreation. Because of this, results of our study should not be evaluated in the context of causality. Rather, we believe the study is offering valuable information on what kind of changes and differences could be observed during a long-lasting intervention. Finally, the workforce was male-dominated. It is not clear if similar results could be obtained in a female-oriented organization.

CONCLUSION

In our study, we found high participation in both HRAs and biometrical screenings (≥80%) without incentives. The annual participation rate into the intervention was modest, but the multiyear continuum helped to accumulate attendance to over 50%. This was

likely due to the multiple entry channels and different services. Approximately one in five employees achieved the a priori defined criterion of success. Our findings indicate that a larger part of the population may report lifestyle changes, but only a smaller part can improve multiple health outcomes or maintain good results in long term. Based on the PIPE Impact Metric evaluations, the program achieved an 18% net impact in 2010 to 2013, and 14% in 2014 to 2017, and these can now be considered as benchmark values for a multiyear WHP program. However, our results may underestimate total health behavior change and overall health impact achievements due to the set criteria for participation and success. Our findings agree, however, with earlier reviews that suggest that a comprehensive WHP program can promote lifestyle changes and health improvements.^{31–35} Further investigation is needed to clarify the possible health risk changes during the program, and financial perspectives should be incorporated into the future research.

ACKNOWLEDGMENTS

The authors thank Jouni Lahti, PhD, Adjunct Professor, Department of Public Health, University of Helsinki, for double-checking the collected data and the participation and effectiveness analyses.

REFERENCES

1. Shain M, Kramer DM. Health promotion in the workplace: framing the concept; reviewing the evidence. *J Occup Environ Med.* 2004;61:643–648.
2. Pronk N. Best practice design principles of worksite health and wellness programs. *ACSM Health Fitness J.* 2014;18:42–46.
3. Pronk NP. Fitness of the US workforce. *Annu Rev Public Health.* 2015;36:131–149.
4. O’Donnell M. *Health Promotion in the Workplace.* 5th ed, Art & Science of Health Promotion Institute; 2017.
5. Goetzel RZ, Henke RM, Tabrizi M, et al. Do workplace health promotion (wellness) programs work? *J Occup Environ Med.* 2014;56:927–934.
6. Goetzel RZ, Ozminkowski RJ. The health and cost benefits of work site health-promotion programs. *Ann Rev Public Health.* 2008;29:303–323.
7. Goetzel RZ, Shechter D, Ozminkowski RJ, Marmet PF, Tabrizi MJ, Roemer EC. Promising practices in employer health and productivity management efforts: findings from a benchmarking study. *J Occup Environ Med.* 2007;49:111–130.
8. Robroek SJ, van Lenthe FJ, van Empelen P, Burdorf A. Determinants of participation in worksite health promotion programmes: a systematic review. *Int J Behav Nutr Phys Act.* 2009;6:26–126.
9. Towers Watson. *Staying@WorkTM Survey Report, “Pathway to Health and Productivity.”* TW-NA-2011-19283, 2011/2012.
10. Bull S, Gillette C, Glasgow R, Estabrooks P. Worksite health promotion research: to what extent can we generalize the results and what is needed to translate research to practice? *Health Educ Behav.* 2003;30:537–549.
11. Merrill R, Aldana S, Garrett J, Ross C. Effectiveness of a workplace wellness program for maintaining health and promoting healthy behaviors. *J Occup Environ Med.* 2011;53:782–787.
12. Henke R, Goetzel R, McHugh J, Isaac F. Recent experience in health promotion at Johnson & Johnson: lower health spending, strong return of investment. *Health Aff.* 2011;30:490–499.
13. Glasgow R, McCaul K, Fisher K. Participation in worksite health promotion: a critique of the literature and recommendations for future practice. *Health Educ Q.* 1993;20:391–408.
14. Soler R, Leeks K, Razi S, et al. A systematic review of selected interventions for worksite health promotion. The assessment of health risks with feedback. *Am J Prev Med.* 2010;38(2 Suppl.):S237–S262.
15. Taitel M, Haufle V, Heck D, et al. Incentives and other factors associated with employee participation in health risk assessment. *J Occup Environ Med.* 2008;50:863–872.
16. Seaverson E, Grossmeier J, Miller T, Anderson D. The role of incentive design, incentive, value, communications strategy, and worksite culture on health risk assessment participation. *Am J Health Promot.* 2009;23:343–352.
17. Marshall AL. Challenges and opportunities for promoting physical activity in the workplace. *J Sci Med Sport.* 2004;7(1 Suppl.):60–66.
18. Loeppke R, Edington D, Bég S. Impact of the prevention plan on employee health risk reduction. *Popul Health Manag.* 2010;13:275–284.

19. Serxner S, Anderson D, Gold D. Building program participation: strategies for recruitment and retention in worksite health promotion programs. *Am J Health Promot.* 2004;18:1–6. iii.
20. Yen L, Schultz AB, Schaefer C, Bloomberg S, Edington DW. Longterm return on investment of an employee health enhancement program at a Midwest utility company from 1999 to 2007. *Int J Workplace Health Manag.* 2010;3:79–96.
21. Byrne D, Goetzel R, McGown P, et al. Seven-year trends in employee health habits from a comprehensive workplace health promotion program at Vanderbilt University. *J Occup Environ Med.* 2011;53:1372–1381.
22. Goetzel R, Guindon A, Turshen J, Ozminkowski R. Health and productivity management: establishing key performance measures, benchmarks, and best practices. *J Occup Environ Med.* 2001;43:10–17.
23. Mercer 2018 North America Mercer Turnover Survey—website. Available at: <https://www.imercer.com/content/article/employee-turnover.aspx>. Accessed May 29, 2019.
24. Rongen A, Robroek S, Ginkel W, Lindeboom D, Altink B, Burdorf A. Barriers and facilitators for participation in health promotion programs among employees: a six-month follow-up study. *BMC Public Health.* 2014;14:573.
25. Middlestadt S, Sheats J, Geshnizjani A, Sullivan M, Arvin C. Factors associated with participation in work-site wellness programs: Implications for increasing willingness among rural service employees. *Health Educ Behav.* 2011;38:502–509.
26. Hall J, Kelly K, Burmeister L, Merchant J. Workforce characteristics and attitudes regarding participation in Worksite Wellness programs. *Am J Health Promot.* 2017;31:391–400.
27. Kilpatrick M, Blizzard L, Sanderson K, Teale B, Venn A. Factors associated with availability of, and employee participation in comprehensive workplace health promotion in a large and diverse Australian public sector setting: a cross-sectional survey. *J Occup Environ Med.* 2015;57:1197–1206.
28. Batorsky B, Stolk C, Liu H. Is more always better in designing workplace wellness programs? A comparison of wellness program components versus outcomes. *J Occup Environ Med.* 2016;58:987–993.
29. Hammer S, Liebherr M, Kersten S, Haas C. Adherence to worksite health interventions: practical recommendations based on a theoretical framework. *J Workplace Behav Health.* 2015;30:325–343.
30. Lerner D, Rodday A, Cohen J, Rogers W. A systematic review of the evidence concerning the economic impact of employee-focused health promotion and wellness programs. *J Occup Environ Med.* 2013;55:209–222.
31. Malik S, Blake H, Suggs S. A systematic review of workplace health promotion interventions for increasing physical activity. *Brit J Health Psych.* 2014;19:149–180.
32. Conn V, Hafdahl A, Cooper P, Brown L, Lusk S. Meta-analysis of workplace physical activity interventions. *Am J Prev Med.* 2009;37:330–339.
33. Osilla K, Van Busum K, Schnyer C, et al. Systematic review of the impact of worksite wellness programs. *Am J Manage Care.* 2012;18:e68–e81.
34. Rongen A, Robroek S, van Lenthe F, Burdorf A. Workplace health promotion: a meta-analysis of effectiveness. *Am J Prev Med.* 2013;44:406–415.
35. Baicker K, Cutler D, Song Z. Workplace wellness programs can generate savings. *Health Aff.* 2010;29:304–311.
36. Pronk N. Designing and evaluating health promotion programs: simple rules for a complex issue. *Dis Manage Health Outcomes.* 2003;11:149–157.
37. Äikäs A, Pronk N, Hirvensalo M, Absetz P. Does implementation follow design? A Case Study of a Workplace Health Promotion Program Using the 4-S Program Design and the PIPE Impact Metric Evaluation Models. *J Occup Environ Med.* 2017;8:752–760.
38. Ovretveit J. *Evaluating Improvement and Implementation for Health.* New York: Open University Press, McGraw Hill Education; 2014.
39. Estabrooks P, Brownson R, Pronk N. Dissemination and implementation science for public health professionals: an overview and call to action. *Prev Chronic Dis.* 2018;20:E162.
40. Glasgow R, Goldstein M, Ockene J, Pronk N. Translating what we have learned into practice: principles and hypotheses for addressing multiple behaviors in primary care. *Am J Prev Med.* 2004;27:88–101.
41. Pronk N. Practice and research connected: a synergistic process of research translation through knowledge transfer. In: Pronk NP, editor. *ACSM’s Worksite Health Handbook.* 2nd ed, Champaign, IL: Human Kinetics; 2009. A Guide to Building Healthy and Productive Companies. Chapter 11.
42. Aziz Z, Absetz P, Oldroyd J, Pronk N, Oldenburg B. A systematic review of real-world diabetes prevention programs: learning from the last 15 years. *Implement Sci.* 2015;10:172.

43. Aziz Z, Riddell M, Absetz P, Brand M, Oldenburg B. Peer support to improve diabetes care: an implementation evaluation of the Australasian Peers for Progress Diabetes Program. *BMC Public Health*. 2018;18:262.
44. O'Reilly S, Dunbar J, Versace V, et al. Mothers after gestational diabetes in Australia (MAGDA): a randomised controlled trial of a postnatal diabetes prevention program. *PLoS Med*. 2016;13:1–21.
45. Lindberg R. Active living: on the road with the 10,000 steps program. *J Am Diet Assoc*. 2000;100:878–879.
46. Linnan L, Bowling M, Childress J, et al. Results of the 2004 national worksite health promotion survey. *Am J Public Health*. 2008;98:1503–1509.
47. Prochaska J, Velicer W. The transtheoretical model of health behavior change. *Am J Health Promot*. 1997;12:38–48.
48. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. *Health Psychol*. 2008;27:379–387.
49. Laukkanen R, Collins J, Scroderus J. *Polar Body Age: Scientific Background and Content in Polar OwnTestTMSsystem*. Oulu, Finland: Polar Electro Ltd; 2011, White paper.
50. Polar Electro Ltd. Polar BodyAge HRA assessment. Available at: https://support.polar.com/e_manuals/BodyAge/Polar_BodyAge_user_manual_English.pdf. Accessed May 29, 2019.
51. 4event Ltd. Baseline HRA report for Stora Enso Metsä Corporation. Lahti, Finland. 2010.
52. 4event Ltd. HRA questionnaire for Stora Enso Metsä Corporation 2010–2017. Lahti, Finland. 2010.
53. Helakorpi S, Pajunen T, Jallinoja P, Virtanen S, Uutela A. National Institute for Health and Welfare Report 15. Health Behaviour and Health Among the Finnish Adult Population, Spring 2010. Helsinki, Finland; 2011.
54. Liukkonen M, Nygård C, Laukkanen R. A cluster randomized controlled trial on the effects of technology-aided testing and feedback on physical activity and biological age among employees in a medium-sized enterprise. *Saf Health Work*. 2017;8:393–397.
55. University of Jyväskylä. Ethical Principles of University of Jyväskylä. Available at: <https://www.jyu.fi/hallinto/strategia/en/ethical-principles>. Accessed November 10, 2018.
56. Gemson D, Commisso R, Fuente J, Newman J, Benson S. Promoting weight loss and blood pressure control at work: impact of an education and intervention program. *J Occup Environ Med*. 2008;50:272–281.
57. Matke S, Liu H, Caloyeras J, et al. *Workplace Wellness Program Study: Final Report*. Santa Monica, CA: RAND; 2013.
58. Racette S, Deusinger S, Inman C, et al. Worksite Opportunities for Wellness (WOW): effects on cardiovascular disease risk factors after 1 year. *Prev Med*. 2009;49:108–114.
59. Elliot D, Goldberg L, Duncan T, et al. The PHLAME firefighters' study: feasibility and findings. *Am J Health Behav*. 2004;28:13–23.
60. MacKinnon D, Elliot D, Thoemmes F, et al. Long-term effects of a worksite health promotion program for firefighters. *Am J Health Behav*. 2010;34:695–706.
61. Harden S, Gaglio B, Shoup J, et al. Fidelity to and comparative results across behavioral interventions evaluated through the RE-AIM framework: a systematic review. *Syst Rev*. 2015;4:155.
62. Edington DW. Emerging research: a view from one research center. *Am J Health Promot*. 2001;15:341–349.