Effectiveness of a mHealth Coaching Program on Predictors of Work Absenteeism

Bojan Simoski(✉) and Michel C. A. Klein

Department of Computer Science, VU Amsterdam, Amsterdam, The Netherlands
{b.simoski,michel.klein}@vu.nl

Abstract. Health-related work absenteeism has an extensive economic and societal impact. Digital health interventions at the workplace can play a beneficial role in reducing the risks of absenteeism. However, the effects of these health interventions are rarely explored in terms of predicting work absenteeism. This paper presents the outcomes of a six month coaching-based digital health intervention. In this intervention, employees receive health or lifestyle coaching by using a smartphone app. We define eight predictors of absenteeism based on an extensive validated health check that the participants have filled throughout the program participation. The predictors are related to mental health, work ability, work stress, physical activity, perceived health and need to recovery after work. We show statistically significant effects of change in multiple predictors of absenteeism as a result of being part of the intervention. Additionally, we define multiple app usage scores founded on the coaching-based smartphone app. They are related to frequency of coach-coachee communication or doing healthy activities. We correlate these app scores with the predictors of absenteeism scores, and detect moderately-strong and significant correlations.

Keywords: Mobile health · Workplace digital health intervention · Coaching program

1 Introduction

The estimated economic and societal impact of health related absenteeism is enormous. For example, estimates of the costs of absence in the UK range from £270 to £659 per employee per year [8]. In an estimation of the costs in the US, the economic burden of illness was assessed up to $392 per employee per year [10]. The costs of just back pain to society in the Netherlands in 1991 were assessed to be 1.7% of the GNP [25]. Interventions focusing on improving health and lifestyle via employers are therefore attractive. A meta-analysis of workplace disease prevention and wellness programs shows that such interventions have both economic and health benefits [1]. In their critical meta-analysis of the

X. Gao et al. (Eds.): MobiHealth 2021, LNICST 440, pp. 204–220, 2022.
https://doi.org/10.1007/978-3-031-06368-8_14

literature on costs and savings associated with disease prevention and wellness programs to improve health, it was found that absenteeism costs fall by about $2.73 for every dollar spent on these programs. Programs focusing on physical activity (PA) have also reported positive results. For example, [12] found that an individually tailored intelligent physical exercise training with recommendations of leisure-time PA significantly decreased absenteeism when following the protocol. Participation in an employee fitness program resulted in a significant decline in sick days (4.8 days) for people in the high participation group [14]. However, a systematic review on the effectiveness of physical activity programs at worksites reported limited evidence of an effect on absenteeism [19].

In this paper, we investigated the effect of a mhealth coaching program in which employees receive health coaching from lifestyle coaches by using an smartphone app. VitalityPlatform (anonymous industry partner) is a platform that combines software with human coaching, where a lifestyle coach professional - a coach, communicates with an employee participating in the program - a coachee, via a smartphone app. This paper analyses the changes on different predictors of absenteeism (measured via questionnaires) through the participation in the program. We investigate the effects of the health intervention on these absenteeism predictors. In addition, we define a set of VitalityPlatform app usage measurements related to the frequency of coach-coachee communication, and the activities performed by the coachee. We investigate the relationship between these app measurements and the predictors of absenteeism scores.

This paper is structured as follows. In Sect. 2 we present the related work on predictors of sick leave and absenteeism, followed by an overview of coaching-based digital health interventions in the workplace. The methodology (Sect. 3) gives more details on the health intervention program, the predictors of absenteeism scores and VitalityPlatform app scores, and the data analysis methods. Section 4 shows the results of the study related to the effect of the intervention on the predictor scores, and the relationship between the app measurements and these scores. These outcomes are discussed in the final Sect. 5.

2 Related Work

2.1 Predictors of Sick Leave and Absenteeism

A literature study has been performed to investigate predictors of sick leave and absenteeism. In [18], the effect of PA on sick leave is investigated. No differences was found in mean sick leave duration between those who met the moderate intensity recommendation and those who did not. However, a significant reduction in amount of sick leave in workers meeting the vigorous intensity activity recommendation was found, as well as a dose-response relation between frequency of vigorous intensity activity up to a frequency of 3 times a week and sick leave duration. Workers not carrying out vigorous activity at all had the most days of sick leave, whereas those who were vigorously active three times a week had the least sick leave. PA was measured using the SQUASH questionnaire [26].

In [2], it is investigated which factors can improve work participation for people with a chronic illness. The effects of perceived health and limitations at work were investigated. The conclusion is that the association between health and sick leave can be explained by limitations at work, work characteristics, and work adjustments. Perceived health was assessed with a single question: 'how do you evaluate your health in general?'.

In [7], the effectiveness of a worksite social and physical environment intervention on need for recovery (i.e., early symptoms of work-related mental and physical fatigue), PA and relaxation was investigated. The "need for recovery after work" is mentioned as an early indicator for mentally and physically work-induced fatigue. Also, PA helps unwinding from work and reduces levels of stress. Measurements that are used are the SQUASH [26] for physical activity, the Oldenburg Burnout Inventory [20], the Need for recovery questionnaire [9], and a generic question about the perceived health. In [21], depression and low self-rated health are identified as risk factors for longer sick-leave periods. Measurements that are used are the ability to work [11] and the PHQ-9 [13].

2.2 Health Coaching Programs in the Workplace

Workplace interventions to improve health are popular for several reasons. People spend much time at work, there are existing social structures that can be used in the interventions, and it is possible to combine them with incentives from the employer [38]. Health interventions at the workplace may have different aims. Many of them focus on increasing PA. A meta-review by [27] has identified more than 138 reports describing studies on workplace PA interventions, showing that the interventions are very diverse and some can improve health. Another review [28] also shows a growing evidence base that workplace PA-based interventions can positively influence PA behaviour. Another, more recent focus for workplace interventions is breaking prolonged sedentary behavior. A systematic review of 26 studies aiming at reducing sitting among white-collar working adults concludes that there is significant overall effect on workplace sitting reduction [34]. A review of 25 interventions using mobile technology concludes that there is reasonable evidence for mHealth in a workplace context as a feasible, acceptable and effective tool to promote PA and reduce sitting time [35].

Mental health is another target of workplace interventions. Many interventions use mindfulness [29] to reduce stress at the workplace [30], others apply psycho-education [31] or psycho-social approaches [32]. A recent review has found evidence for the effectiveness of workplace interventions on the prevention of mental health problems [33].

Several workplace interventions focus on multiple health goals, e.g. a combination of dietary behaviour and PA [36,37]. Overall, health coaching to improve lifestyle behaviors seems to be a promising way to prevent diseases. In a 2010 literature review, significant improvements in one or more of the behaviors of nutrition, physical activity, weight management, or medication adherence were found in six (40%) of the fifteenth studies [17]. A more recent review of 18 interventions found small positive effect sizes, which were partly determined by intervention characteristics and experimental setup [38].

A few studies focus specifically on coaching as the means of the intervention. Coaching is about enhancing well-being and performance in personal life and work domains with normal, non-clinical populations, underpinned by models of coaching grounded in established adult learning or psychological approaches [39]. A recent meta-analysis has shown a positive effect of coaching in organizations. Interestingly, no difference in effect was found for different coaching formats (comparing face-to-face, with blended face-to-face and e-coaching) or duration of coaching (number of sessions or longevity of intervention). A study about the effect of coaching on the reduction of workplace stress has shown mixed outcomes: anxiety and stress had decreased more in the coaching group compared to the control group, but levels of depression had decreased more in the control group compared to the coaching group. However, participants reported high levels of perceived coaching effectiveness [40].

Overall we can conclude that there is some evidence that health coaching, also in a blended from, could have positive effects on the factors related to physical and mental health. In the remainder of this paper, we describe the methods and results of our study on the effectiveness of a specific online health coaching program on predictors of absenteeism.

3 Methodology

3.1 Recruitment and Timeline

Employees from 10 different companies were recruited to potentially participate in the program. Companies were offered a fixed number of places for their employees in the program. Potential participants received an email with an invitation to perform a "vitality scan", an online pre-scan on motivation and potential for improvement. This scan assessed their motivation for increasing physical activity using the SRQ-E [4]. In addition, their current health behavior was assessed by measuring physical activity (using the SQUASH questionnaire [26]), sleep quality (using the PSQ [5]), eating habits (using the RIVM VCP questionnaire [3]) and a work-related burnout [6,23]. All subscales for health behavior have been normalized to a scale between −2,5 and +2,5, where a score of −2,5 represent extremely unhealthy behaviors and a score of +2,5 represent the most healthy behaviors. The total score for health behavior was calculated by summing up all subscores. The exact questions and the normalization used are available per request.

Based on the outcome of the vitality check, a subset of employees were invited to join the health coaching program. The aim was to include people that were motivated to increase their health behavior and for whom there was some room for improvement. This has been operationalized by inviting people that scored above 0 on the motivation scale and below 1 on the health behavior scale, as shown in Fig. 1.

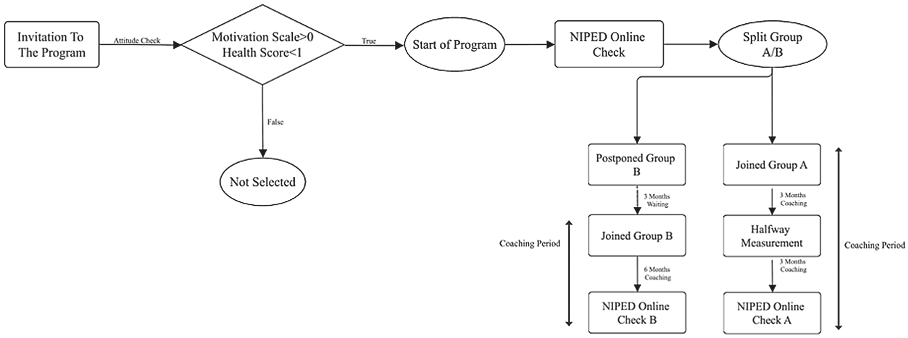

Fig. 1. VitalityPlatform Recruitment flowchart

The selected participants were then randomly assigned and divided in two groups. Group A participants have started the coaching program directly, while group B participants had to wait for three months (control group).

Both groups were invited to fill in a baseline questionnaire, the NIPED personal health check (time0). When the number of places in the program was not filled up by people in group A after two weeks, people were randomly selected from group B to switch to group A. Three months after the baseline measurements, group A participants were invited for a halfway-program NIPED check, while group B started their coaching program (time1). From this time point, group A participants took three more months to finish the program (time2), while group B participants, having started later, finished the program and filled their final NIPED survey after six months (time3).

3.2 NIPED

The NIPED health check [15,16,24] is a modular health check consisting of questionnaires and physical tests, which has been used by more than 325.000 people. The check is often offered by employers or insurance companies. The version used in this research contains 92 questions of items related to mental health, work ability, work stress, nutrition, physical activity, perceived health and the need for recovery, mostly based on existing scales. These questions were used to calculate scores for predictors of sick leave. Table 1 shows the 8 predictors of sick leave (absenteeism), most of them being based on validated questionnaires. The workability predictor is sourced from WAI [11], burnout is based on the UBOS [23], depressive symptoms are derived from the PHQ-2 questionnaire [41], perceived health as used in WHO survey [22], and the need for recovery based on NFR [9]. The last three predictors are based on a single questions about PA. PA/w refers to number of PA minutes per week, VPA/w refers to the number of minutes per week in which participants performed moderate/vigorous PA activities. The dPA30 question asked participants about the number of days in a week in which they move more than 30 min (referring to moderate and vigorous exercises).

Table 1. Predictors of sick leave.

Predictor	Composite score	Final score
Workability	Nine questions, combination of choice-questions and 5-points likert scale	0 (very bad) - 50 (very good)
Burnout	Five questions, 7-points likert scale	0 (very good) - 6 (very bad)
Depression	Two questions, 4-points scale (not at all - several days - more than half the days - nearly every day)	0 (never depressive symptoms) - 6 (nearly every day depressive feelings and no pleasure in doing things)
Perceived Health	One question, 5-points likert scale	1 (very bad) - 5 (very good)
Need for recovery	Eleven questions, yes/no	0 (very good) - 100 (very bad)
PA/w	One question	Number of minutes
VPA/w	One question	Number of minutes
dPA30	One question	Number of days, 0 to 7

3.3 VitalityPlatform Health Coaching Platform

The participants have installed the app that is part of the VitalityPlatform health coaching platform on their smartphones. The app offers mobile coaching experience, where an employee (coachee) can pick their own professional coach. The platform first gives a recommendation for a set of coaches, based on the requirements of coachees for certain lifestyle improvement. If needed, the coachee can anytime change the coach. One coach can have multiple coachees under their guidance. The communication is done via the app, by using chat messages or doing face-to-face video calls. The coachee can inform their coach when a certain exercise has been completed and connect the app to a fitness tracker. In this way, the coach is able to track the daily progress. The coach will continuously update the coachee' activity schedule based on their mutual communication. The coachee can undertake activities divided in several categories: mental, workout, lifestyle, physical and nutrition.

Only the meta-data of the VitalityPlatform usage is available for analysis. Data items are related to the frequency of coach-coachee (dyad) communication: number_messages, number_conversations, number_calls, duration_calls; and the coachee activities: ratio_done, number_done, number_liked. Table 2 gives a more detailed description of these variables.

3.4 Data Analysis

A Wilcoxon signed-rank was applied to the predictors of absenteeism scores at different time points for both groups. We test the effect of change in the predictors of absenteeism as a result of using the app (6 months period, for both groups). Wilcoxon signed-rank test was used as a non-parametric univariate test, applicable when the data violates the assumption of normality, which is the case for the predictors of absenteeism data as we will show in the next chapter.

Table 2. Variables derived from the VitalityPlatform app.

Variable name	Variable description	Possible values
number_messages	Total number of dyad text messages	[0, inf+]
number_conversations	Total number of dyad conversations	[0, inf+]
number_calls	Total number of dyad video/audio calls	[0, inf+]
avg_duration_calls	Average minutes of dyad video/audio calls	[0, inf+]
ratio_done	Ratio of planned versus done activities	[0, 1]
number_done	Total number of done activities	[0, inf+]
number_liked	Total number of liked activities	[0, inf+]

Spearman's rank-order correlation method was used in order to report the relationship between the app measurements and the predictors of absenteeism. For this analysis, the difference in the predictors of absenteeism scores before and after using the VitalityPlatform app, was used as a measure of change.

4 Results

4.1 Descriptive Statistics

Program Participation. In total 560 participants met the selection criteria (for motivation and health check scores) and were thus invited to participate in the coaching program. Figure 2 gives an overview of the number of participants per group, at different time points of the program, based on completing the NIPED questionnaires.

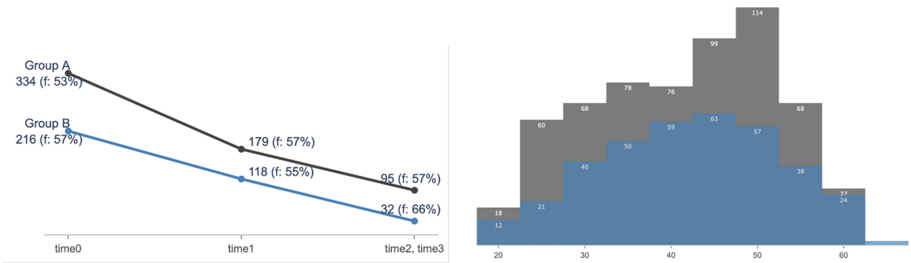

Fig. 2. Completed Niped measurements - program participation. Age distribution.

A group of 334 participants was invited to start directly after taking the baseline survey - group A, another group of 216 participants was offered to start the coaching in 3 months after filling in the baseline survey - group B. As can be seen in Fig. 2, there is a significant dropout in the number of participants who filled the second NIPED questionnaire (time1). At the final point, after completion of group A and group B participation programs (at time points time2 and time3, respectively), we observe another wave of dropout. In total, 95 participants of group A and 31 participants of group B filled the final round of the survey. For the upcoming analysis we consider 93 participants (group A) and 31 participants (group B), as they have completed the three NIPED questionnaires.

Predictors of Absenteeism Scores. Eight predictors of absenteeism were derived based on the NIPED questionnaire, as shown in Table 1. These scores were obtained at all multiple time points of the program for both group A and group B. Below we present an exploratory analysis obtained at the baseline questionnaire (time0), summarized for group A and group B participants (total n = 550, average age = 43.2, 60% female). This analysis was conducted for reporting purposes, and the outcomes were used in order to decide on the relevant statistical tests for answering our research questions.

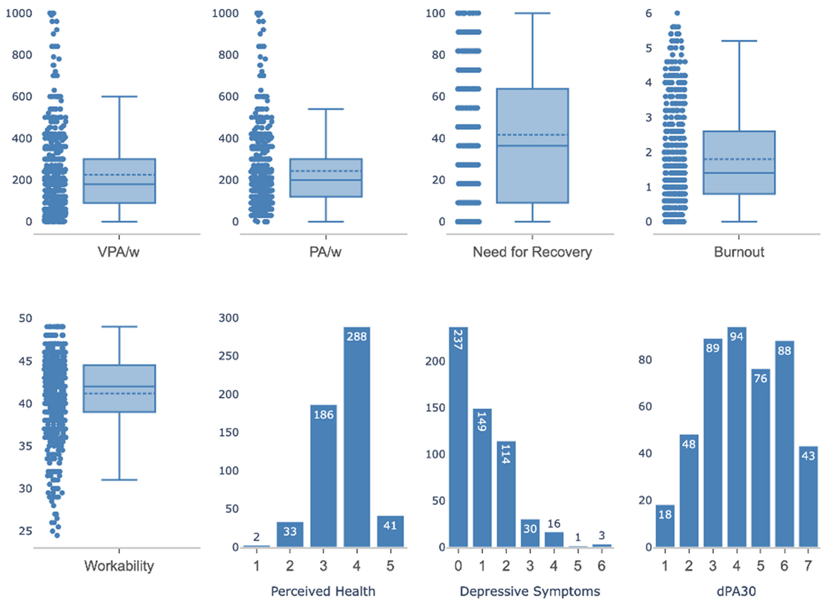

Fig. 3. Data distribution - Predictors of absenteeism scores

Figure 3 visualizes the data distribution of the predictors of absenteeism scores. Five scores are considered as continuous ratio variables, namely: VPA/w, PA/w, need for recovery, burnout and workability scores. The remaining three scores are classified as categorical nominal variables: perceived health, depressive symptoms and dPA30. The participants have self-reported an average of 243.7 min (s = 181.1) of physical activity per week, and an average of 224.9 (s = 194.7) moderate to vigorous minutes activity per week. The average need for recovery after work score was 41.7 (s = 30.78) where a score of 100 is the least favorable score (i.e. a very high need for recovery). The participants have reported a mean score of 1.79 (s = 1.37) for the risk of burnout, where 6 indicates the highest risk of burnout and 0 is the lowest risk. The average workability (or work capacity) score was 41.16 (s = 4.83), with 50 indicating the highest work capacity. Looking at the perceived health score, we can observe that most commonly participants have answered 4-good (n = 288) and 3- not bad, not good (n = 186) with an average response of 3.6 (s = 0.72). Regarding the depression score, participants have dominantly reported 0 - never depressive symptoms (n = 237), followed by 1 - several days either depressive feelings or no pleasure in doing things (n = 149), followed by 2 (n = 114) and 3 (n = 30), 4 (n = 16), 5 (n = 1) and three participants with a score of 6 (almost every day depressive feelings and no pleasure in doing things). The average is 1.00 (s = 1.13). Finally, participants were asked how many days in a week they do at least 30 min of physical activity (score dPA30). The bottom right plot of Fig. 3 shows a relatively uniform distribution with 3-, 4-, 5- and 6- days of more than 30 min activities.

Next we have looked at the normality of the data, as an important factor for selecting the appropriate statistical methods in the subsequent analysis. Both histograms and quantile-quantile (QQ) plots have indicated that all the scores do not follow a normal data distribution. Additionally we have conducted three statistical tests of normality (shapiro-wilk, d'agostino's K2, anderson-darling test) none of which showed a significant p-value regardless of the score. Therefore the considered variables are non-normally distributed. In this case, non-parametric statistical methods are a more suitable solution for analysis. As a result, we have selected the Wilcoxon signed-rank test and Spearman's rank-order method for the analysis presented in the next subsections.

App Usage Scores. This analysis is based on the 124 participants (93- group A, 31- group B) that have used the VitalityPlatform app during a period of six months (time0, time2- group A, time1, time3 - group B) and have completed the NIPED questionnaires needed for our analysis. These participants were trained by 25 professional coaches. Each coach can have multiple coachees, in this sample the number of coachees per coach varies between 1 and 12 (avg = 5). Most of the participants have matched with a coach in July 2019 (65 participants), followed by October 2019 (37 participants). Fewer have started in August and September (10 and 3 participants), or November and December 2019 (3, 4 participants). Five participants got their coach in January 2020.

Figure 4 shows the data distribution of the seven app usage scores. These scores belong to two categories: dyad communication and activity trends. The communication scores are related to number of exchanged messages, number of conversations, number of calls (audio/video) and the duration of the calls (audio/video). The activity scores are related to the coachee performance during the app usage. We look into the number of done activities, the number of liked activities and the ratio of done versus planned activities. All the scores are considered as continuous ratio variables. The dyads have exchanged on average

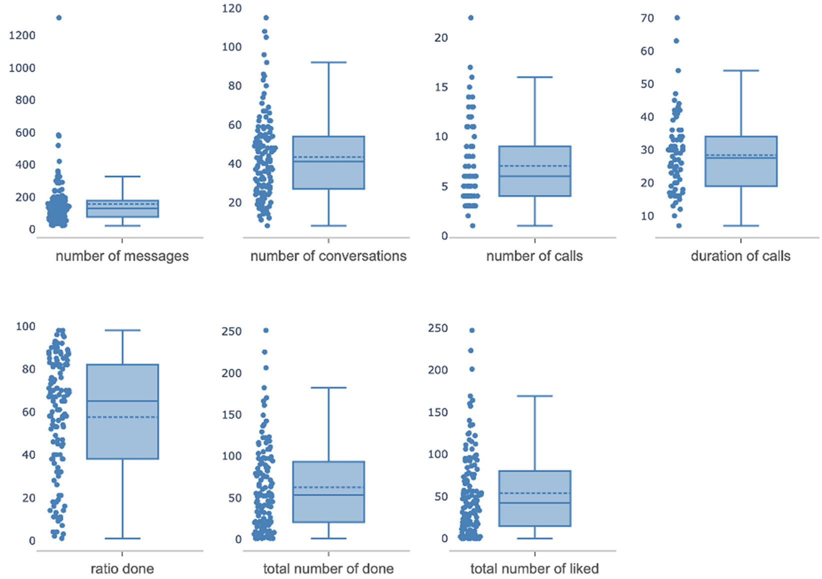

Fig. 4. App measurements data distribution

155 messages (s = 146.50) while using the app. A conversation is defined as any day in which there was a message exchanged between the dyad. On average there were 43 conversations between dyads (s = 21.41). The dyads have communicated via video/audio calls on average 7 times during the app usage (s = 4.2), with an average call time of 28 min (s = 12). Only 70 coachees have used the app for video/audio calls. Moving to the activities, we observe that the coachees did on average 58% (s = 27) of the planned activities. Looking at the total number of finished exercises, the participants did 62 activities on average while using the app (s = 51), and have liked 54 activities (s = 50) out of the ones done on average. In total there were 7728 activities performed via the app. Among them, the most common categories were lifestyle (n = 2128), mental health (n = 2184) and nutrition (n = 2012), followed by physical (n = 1365) and workout (n = 39). The activities were 94% initiated by the coach.

4.2 Effects of the Intervention

Intervention Effect on Predictors of Sick Leave. For analyzing the effects of the intervention on the predictors of sick leave, the absenteeism scores at the end of the program (time2 - group A, time3 - group B) were subtracted from the ones at the beginning of using the app (time 0 - group A, time 1 - group B). These scores are referred to as 'absenteeism difference scores' in the remainder of the paper. They give an indication of potential improvement or deterioration in absenteeism as a result of intervention participation and using the VitalityPlatform app.

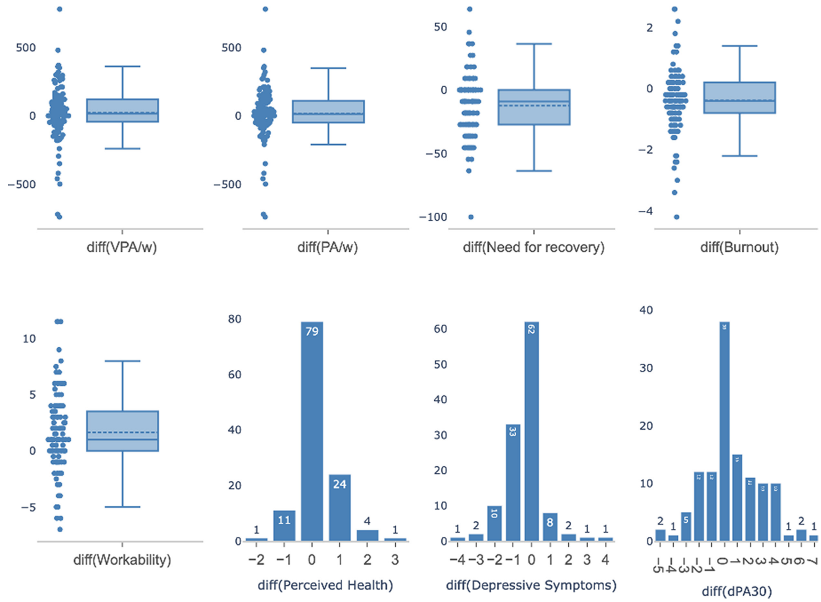

Fig. 5. Data Distribution - absenteeism difference scores

Figure 5 gives a visual overview of the absenteeism difference scores. The depicted results are summarized for group A (n = 93) and group B (n = 31). The participants have reported on average 22.7 min (s = 200.7) more vigorous activities per week - diff(VPA/w), at the end of the program participation. In a related manner, they have reported on average 16.5 more minutes (s = 190) of weekly activities. The reported need for recovery (measured on a scale 0–100, 100 - very bad) was lowered by an average of 12.4 points (s = 24.03). The participants have assessed on average 0.38 (s = 1.03) lower risk of burnout, and reported increased workability of 1.64 (s = 3.43). The five reported absenteeism difference scores are considered as interval continuous variables. The next three scores are categorical variables. First, looking at the diff(Perceived Health) we observe that most of the participants have reported no change (64%), followed by 19% that improved their score by 1, and 3% improved by two. On the other hand, 8% of participants have reduced their score by 1. The depressive symptoms score also sees improvement. While 50% of the participants have no change, 27% have improved their score by 1 (in this case the lower the better), followed by 8% improving the score by 2. One score of deterioration was reported by 6% of the participants. Finally, the diff(dPA30) score shows that 30% of participants have been doing activities in the same number of days of the week. One more day a week of activity reported 12% of participants, followed by 9% reporting doing activities in two more days. Another 8% have seen an improvement of doing three and four more days of activity after finishing the program. Looking at the opposite side, 10% of the participants reported reducing the number of days of activity by one and two days, while 4% did three less days per week.

Table 3. Wilcoxon tests results

	Workability	Burnout	Depressive symptoms	Perceived health	Need for recovery	PA/w	VPA/w	dPA30
Group A	881.5*	899.5*	300.5*	369.0	451.0*	1420.5	1361.5	679.0*
Group B	75.5*	166.5	20.0	8.0*	72.5*	167.0	143.5	95.0

Next we report the effects of change in the predictors of absenteeism as a result of the 6 month intervention for both groups. Table 3 presents the outcomes of running the Wilcoxon tests, with a T statistics and a star mark (*) for statistical significance results. The T statistics is a float value representing the sum of the ranks of the differences above or below zero, whichever is smaller. In Table 3, each T score is associated with a * mark in case of a statistically significant difference between the two means, indicated when p-value < 0.05. We note statistically significant mean differences among five scores for group A participants, namely: workability, burnout, depressive symptoms, need for recovery and dPA30. For group B, we found a statistically significant effect at three scores: workability, perceived health and need for recovery. These results indicate statistically significant changes in the corresponding predictor scores comparing before and after the intervention.

Effects of App Usage Patterns on Predictors of Sick Leave. This analysis is conducted to answer our research question if changes in NIPED scores can be associated with certain app usage patterns, regarding coach-coachee communication or the frequency of activities. We assess the relationship between the app measurements scores and the absenteeism difference scores. The relationships are examined by performing Spearman rank correlation coefficient tests for two reasons: the data is not normally distributed (to use Pearson test) and the reported type of variables. All the app scores are ratio variables, while the absenteeism difference scores are interval ($n = 5$) and categorical ($n = 3$). The results are presented separately for group A and group B, as reported in Table 4 and Table 5. The statistically significant correlation coefficients (with $p < 0.05$) are marked with a star sign (*). Among group A participants we observe four statistically significant relationships between the app usage measurements and the absenteeism difference scores. The number of dyad audio/video calls is linked with higher PA/w ($r = 0.30$) and need for recovery ($r = 0.49$), the number of exchanged messages has a moderate relation with the increase of perceived health ($r = 0.24$), and the number of liked activities with the workability score ($r = 0.24$). As observed in Table 5, among group B participants, the exchanged number of messages could be related to reducing the risk of burnout ($r = -0.40$), while the ratio of done activities lead to increase of the depressive symptoms ($r = -0.39$).

Table 4. Correlation outcomes - groupA

	avg_dur	num_calls	num_conv	ratio_done	num_mes	total_done	total_liked
diff(Workability)	0.14	0.04	0.10	−0.04	0.07	0.19	0.24*
diff(Burnout)	0.01	−0.02	−0.08	0.04	−0.05	0.05	0.04
diff(Depressive symptoms)	−0.04	−0.14	−0.03	−0.05	−0.05	−0.01	0.05
diff(Perceived Health)	−0.24	0.15	0.12	−0.01	0.24*	0.10	0.08
diff(Need recovery)	−0.04	0.49*	0.07	0.14	0.09	0.02	0.02
diff(PA/w)	−0.11	0.30*	0.06	−0.04	−0.01	−0.02	0.01
diff(VPA/w)	−0.14	0.24	0.08	−0.03	0.03	−0.03	−0.01
diff(dPA30)	−0.14	0.10	0.03	0.07	0.02	−0.04	0.03

Table 5. Correlation outcomes - groupB

	avg_dur	num_calls	num_conv	ratio_done	num_mes	total_done	total_liked
diff(Workability)	−0.32	0.35	0.02	−0.15	0.22	0.10	0.08
diff(Burnout)	0.24	−0.20	−0.36	0.11	−0.40*	−0.13	−0.10
diff(Depressive symptoms)	−0.17	0.05	0.34	0.39*	0.20	0.30	0.23
diff(Perceived Health)	−0.29	0.25	0.12	0.02	0.18	0.14	0.14
diff(Need recovery)	0.35	−0.10	−0.32	−0.03	−0.36	−0.32	−0.28
diff(PA/w)	0.07	−0.34	−0.20	0.17	−0.17	0.01	0.01
diff(VPA/w)	0.15	−0.41	−0.20	0.12	−0.19	−0.05	−0.04
diff(dPA30)	−0.20	0.37	0.26	0.07	0.15	0.04	0.06

5 Discussion and Conclusions

This paper presents the outcomes of a coaching-based digital health intervention and its effects on predictors of absenteeism. This makes our work distinguishable from previous research, as most commonly the intervention effects are measured (linked to) health outcomes such as increasing PA or weight loss. The predictors of absenteeism scores were analyzed after 6 months of coach-based intervention via smartphone app. The summarized results show improvement in all the absenteeism difference scores. The participants have reported lower risk of burnout (0.38 points), showing less depressive symptoms and need for recovery (12.4 points). On the other hand, they have performed more PA activities per week (any activities 16.5 min, vigorous 22.7 min), increased their workability (1.64), improved their perceived health and had slightly more days of PA per week. These results indicate that the mhealth intervention had a positive effect on the participants, if one is to observe the difference in the reported scores before and after the intervention. Furthermore, the outcome of the Wilcoxon tests shows statistically significant effects of change in the predictors of absenteeism as a result of using the app. This has been observed separately in both group A and group B participants. The effects are detected in the following absenteeism scores: workability, burnout, depressive symptoms, need for recovery and dPA30 (group A); perceived health, workability and need for recovery (group B).

Additionally, we have investigated the relationship between the VitalityPlatform app usage (in terms of frequency of dyad communication and performing activities) and the observed changes in the absenteeism predictor scores. The spearman correlation coefficients did not find any strong and significant correlation between the two. However, we are able to detect moderately-strong and significant correlations. Regarding the frequency of dyad communication we found significant relationships between the number of audio/video calls with need for recovery and PA/w (group A), and the number of exchanged messages correlates with the perceived health (group A), and burnout (group B). Looking at the performed activities, the total number of liked activities was linked to the workability score (group A), and surprisingly the depressive symptoms increase with the ratio of done exercises (group B).

This work comes with certain limitations. Only a limited number of people filled in the NIPED questionnaire after three months. Engagement data shows that 90% of the participants were active after 3 months. Therefore, the reason behind the relatively low number of participants who completed the second test was not that they dropped out of the health coaching program. We suspect that that the relative cumbersome health check (it takes 30 to 45 min to complete the entire test) is the primary reason for people not completing the second questionnaire. For a more thorough analysis of the effect of the coaching program on itself, we ideally need a control group data that is not invited to participate in a health coaching program. With the current experimental setup, we cannot explicitly investigate this, as group B participants were put on a waiting list. One possible additional analysis could be whether the participation in the programme or already filling in the first health questionnaire (for group B) is causing a potential positive effect among participants.

The investigated coaching aspect also comes with certain constraints. For example, we cannot rule out that there have been other interactions between the dyad then via the app. We have focused on quantifying the dyad relationship through frequency of communication. However, textual analysis of the chat interactions can offer additional valuable information on their relationship. The defined app usage measurements are of course a simplification. Many other (non) latent factors might potentially influence the effects of the coaching program, for example seasonal effects. Therefore a larger and more complete study would be prefered to investigate the relation between the app usage and change in predictors of absenteeism.

References

1. Baicker, K., Cutler, D., Song, Z.: Workplace wellness programs can generate savings. Health Aff. (2010). https://doi.org/10.1377/hlthaff.2009.0626
2. Boot, C.R.L., Koppes, L.L.J., van den Bossche, S.N.J., Anema, J.R., van der Beek, A.J.: Relation between perceived health and sick leave in employees with a chronic illness. J. Occup. Rehabil. **21**(2), 211–219 (2011). https://doi.org/10.1007/s10926-010-9273-1
3. van den Brink, C., Ocké, M., Houben, A., van Nierop, P., Droomers, M.: Validering van Standaardvraagstelling voeding voor Lokale en Nationale Monitor Volksgezondheid. RIVM Rapport 260854008 (2005)
4. Brown, J.M., Miller, W.R., Lawendowski, L.A.: The Self-Regulation Questionnaire (SRQ). Innovations in Clinical Practice: A Source Book, vol. 17 (1999)

5. Buysse, D.J., Reynolds, C.F., Monk, T.H., Berman, S.R., Kupfer, D.J.: The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. Psychiatry Res. (1989). https://doi.org/10.1016/0165-1781(89)90047-4
6. Champion, D.F., Westbrook, B.W.: Maslach burnout inventory. Meas. Eval. Couns. Dev. (1984). https://doi.org/10.1080/07481756.1984.12022754
7. Coffeng, J.K., Boot, C.R.L., Duijts, S.F.A., Twisk, J.W.R., Van Mechelen, W., Hendriksen, I.J.M.: Effectiveness of a worksite social & physical environment intervention on need for recovery, physical activity and relaxation; results of a randomized controlled trial. PLoS ONE (2014). https://doi.org/10.1371/journal.pone.0114860
8. Cooper, C., Dewe, P.: Well-being-absenteeism, presenteeism, costs and challenges. Occup. Med. **58**(8), 522–524 (2008). https://doi.org/10.1093/occmed/kqn124
9. de Croon, E.M.: Psychometric properties of the Need for Recovery after work scale: test-retest reliability and sensitivity to detect change. Occup. Environ. Med. **63**(3), 202–206 (2006). https://doi.org/10.1136/oem.2004.018275
10. Goetzel, R.Z., Long, S.R., Ozminkowski, R.J., Hawkins, K., Wang, S., Lynch, W.: Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting U.S. employers. J. Occup. Environ. Med. (2004). https://doi.org/10.1097/01.jom.0000121151.40413.bd
11. Ilmarinen, J.: The work ability index (WAI). Occup. Med. **57**(2), 160–160 (2006). https://doi.org/10.1093/occmed/kqm008
12. Justesen, J.B., Søgaard, K., Dalager, T., Christensen, J.R., Sjøgaard, G.: The effect of intelligent physical exercise training on sickness presenteeism and absenteeism among office workers. J. Occup. Environ. Med. (2017). https://doi.org/10.1097/JOM.0000000000001101
13. Kroenke, K., Spitzer, R.L., Williams, J.B.W.: The PHQ-9: validity of a brief depression severity measure. J. Gen. Intern. Med. (2001). https://doi.org/10.1046/j.1525-1497.2001.016009606.x
14. Lechner, L., De Vries, H., Adriaansen, S., Drabbels, L.: Effects of an employee fitness program on reduced absenteeism. J. Occup. Environ. Med. (1997). https://doi.org/10.1097/00043764-199709000-00005
15. Niessen, M.A.J., et al.: Short term reduction in absenteeism after implementation of a personalized prevention program. Eur. J. Cardiovasc. Prev. Rehabil. (2010)
16. Niessen, M.A.J., Kraaijenhagen, R.A., Dijkgraaf, M.G.W., Van Pelt, D., Van Kalken, C.K., Peek, N.: Impact of a web-based worksite health promotion program on absenteeism. J. Occup. Environ. Med. (2012). https://doi.org/10.1097/JOM.0b013e31824d2e43
17. Olsen, J.M., Nesbitt, B.J.: Health coaching to improve healthy lifestyle behaviors: an integrative review. Am. J. Health Promot. **25**(1), e1–e12 (2010)
18. Proper, K.I.: Dose-response relation between physical activity and sick leave. Br. J. Sports Med. **40**(2), 173–178 (2006). https://doi.org/10.1136/bjsm.2005.022327
19. Proper, K.I., Staal, B.J., Hildebrandt, V.H., van der Beek, A.J., van Mechelen, W.: Effectiveness of physical activity programs at worksites with respect to work-related outcomes. Scand. J. Work Environ. Health **28**(2), 75–84 (2002). https://doi.org/10.5271/sjweh.651
20. Reis, D., Xanthopoulou, D., Tsaousis, I.: Measuring job and academic burnout with the Oldenburg Burnout Inventory (OLBI): factorial invariance across samples and countries. Burn. Res. (2015). https://doi.org/10.1016/j.burn.2014.11.001
21. Salomonsson, S., Hedman-Lagerlöf, E., Öst, L.G.: Sickness absence: a systematic review and meta-analysis of psychological treatments for individuals on sick leave

due to common mental disorders. Psychol. Med. (2018). https://doi.org/10.1017/S0033291718000065

22. Subramanian, S.V., Huijts, T., Avendano, M.: Self-reported health assessments in the 2002 World Health Survey: how do they correlate with education? Bull. World Health Organ. **88**(2), 131–138 (2010). https://doi.org/10.2471/BLT.09.067058
23. Schaufeli, W., Van Dierendonck, D.: Utrechtse Burnout Schaal (UBOS). De Psycholoog (2001)
24. Van Den Brekel-Dijkstra, K., Rengers, A.H., Niessen, M.A.J., De Wit, N.J., Kraaijenhagen, R.A.: Personalized prevention approach with use of a web-based cardiovascular risk assessment with tailored lifestyle follow-up in primary care practice - a pilot study. Eur. J. Prev. Cardiol. (2016). https://doi.org/10.1177/2047487315591441
25. van Tulder, M.W., Koes, B.W., Bouter, L.M.: A cost-of-illness study of back pain in The Netherlands. Pain (1995). https://doi.org/10.1016/0304-3959(94)00272-G
26. Wendel-Vos, G.C.W., Schuit, A.J., Saris, W.H.M., Kromhout, D.: Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. J. Clin. Epidemiol. (2003). https://doi.org/10.1016/S0895-4356(03)00220-8
27. Conn, V.S., Hafdahl, A.R., Cooper, P.S., Brown, L.M., Lusk, S.L.: Meta-analysis of workplace physical activity interventions. Am. J. Prev. Med. **37**(4), 330–339 (2009)
28. Dugdill, L., Brettle, A., Hulme, C., McCluskey, S., Long, A.F.: Workplace physical activity interventions: a systematic review. Int. J. Workplace Health Manag. (2008)
29. Chiesa, A., Serretti, A.: Mindfulness-based stress reduction for stress management in healthy people: a review and meta-analysis. J. Alternative Complement. Med. **15**(5), 593–600 (2009)
30. Wolever, R.Q., et al.: Effective and viable mind-body stress reduction in the workplace: a randomized controlled trial. J. Occup. Health Psychol. **17**(2), 246 (2012)
31. Kagan, N.I., Watson, M.G.: Stress reduction in the workplace: the effectiveness of psychoeducational programs. J. Couns. Psychol. **42**(1), 71 (1995)
32. Czabała, C., Charzyńska, K., Mroziak, B.: Psychosocial interventions in workplace mental health promotion: an overview. Health Promotion Int. **26**(Suppl. 1), i70–i84 (2011)
33. Proper, K.I., van Oostrom, S.H.: The effectiveness of workplace health promotion interventions on physical and mental health outcomes-a systematic review of reviews. Scand. J. Work Environ. Health **45**(6), 546–559 (2019)
34. Chu, A.H., Ng, S.H., Tan, C.S., Win, A.M., Koh, D., Müller-Riemenschneider, F.: A systematic review and meta-analysis of workplace intervention strategies to reduce sedentary time in white-collar workers. Obes. Rev. **17**(5), 467–481 (2016)
35. Buckingham, S.A., Williams, A.J., Morrissey, K., Price, L., Harrison, J.: Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: a systematic review. Digit. Health **5**, 2055207619839883 (2019)
36. Hutchinson, A.D., Wilson, C.: Improving nutrition and physical activity in the workplace: a meta-analysis of intervention studies. Health Promot. Int. **27**(2), 238–249 (2012)
37. Verweij, L.M., Coffeng, J., van Mechelen, W., Proper, K.I.: Meta-analyses of workplace physical activity and dietary behaviour interventions on weight outcomes. Obes. Rev. **12**(6), 406–429 (2011)
38. Rongen, A., Robroek, S.J., van Lenthe, F.J., Burdorf, A.: Workplace health promotion: a meta-analysis of effectiveness. Am. J. Prev. Med. **44**(4), 406–415 (2013)

39. Palmer, S., Tubbs, I., Whybrow, A.: Health coaching to facilitate the promotion of healthy behaviour and achievement of health-related goals. Int. J. Health Promot. Educ. **41**(3), 91–93 (2003)
40. Gyllensten, K., Palmer, S.: Can coaching reduce workplace stress? A quasi-experimental study. Int. J. Evidence Based Coaching Mentoring **3**(2), 75–85 (2005)
41. Arroll, B., et al.: Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. Ann. Family Med. **8**(4), 348–353 (2010)