Internet Interventions 37 (2024) 100762

Contents lists available at ScienceDirect

Internet Interventions

journal homepage: www.elsevier.com/locate/invent

Effects of a digital intervention on physical activity in adults: A randomized controlled trial in a large-scale sample

Urs Alexander Fichtner [a,*,1], Iris Tinsel [a,1], Matthias Sehlbrede [a], Phillip Maiwald [a], Martina Bischoff [a], Gloria Metzner [a], Christian Schlett [a], Judith Brame [b], Jan Kohl [b], Daniel König [b,c,d], Rainer Bredenkamp [e], Ramona Wurst [b,2], Erik Farin-Glattacker [a,2]

[a] Section of Health Care Research and Rehabilitation Research, Institute of Medical Biometry and Statistics, Faculty of Medicine, Medical Center—University of Freiburg, 79106 Freiburg, Germany
[b] Department of Sport and Sport Science, University of Freiburg, 79117 Freiburg, Germany
[c] Department of Sport Science, Institute for Nutrition, Exercise and Health, University of Vienna, Vienna, Austria
[d] Department of Nutritional Sciences, Institute for Nutrition, Exercise and Health, University of Vienna, Vienna, Austria
[e] Universitäres Cancer Center Hamburg, Hubertus Wald Tumorzentrum, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany

A R T I C L E I N F O

Keywords:
Web-based program
Health behavior change
Interactive coaching
Physical fitness
Online coaching
Digital health intervention

A B S T R A C T

Background: Physical inactivity is associated with health risks, contributing to various diseases and all-cause mortality. Despite recommendations for regular physical activity (PA), many adults remain inactive, influenced by socioeconomic and environmental factors. Digital interventions, particularly web-based PA programs, offer promising possibilities to promote PA across populations. These programs vary in their effectiveness, reflecting differences in design, user engagement, and behavior change techniques employed.
Objective: This study evaluates the effectiveness of the 12-week multimodal web-based TKFitnessCoach. The PA online program is part of the TK-HealthCoach. This study investigates the program's impact on self-reported PA levels, goal attainment, healthrelated quality of life, body weight, and eating behavior, comparing an interactive personalized web-based intervention and non-interactive web-based health information.
Methods: In a randomized controlled trial (RCT), participants were allocated to either the intervention group (IG), receiving access to the interactive TK-FitnessCoach, or the control group (CG) that was provided a static website with evidence-based information on PA. The study targeted a German-speaking adult population interested in improving health behavior. Data was assessed at T0 (beginning of the study), T1 (postintervention), T2, 6 months, and T3, 12 months follow-ups, focusing on self-reported PA at T3 and on various secondary outcomes.
Results: We achieved equally distributed sociodemographics in both the IG and the CG with a mean age of 42.8 (IG), resp. 43.1 years (CG), and female participants of 76.1 % (IG), resp. 74.7 % (CG). PA at baseline was 277.9 min/week in the IG and 273.3 min/week in the CG. Both, the IG ($n = 1153$ in the Intention-to-treat (ITT) dataset) and CG ($n = 1177$ in the ITT dataset) exhibited significant increases in PA over time (IG(T3-T0) = 72.92 min/week; CG(T3-T0) = 74.12 min/week).
However, the study did not find significant differences in the effectiveness of the interactive TK-FitnessCoach compared to the non-interactive control in terms of improving PA and related health outcomes. The intensity of using the TK-FitnessCoach was not associated with PA.
Conclusions: Both programs were effective in promoting PA among adults, with no significant differences observed between the two RCT groups. This highlights the potential of digital interventions in addressing physical inactivity, suggesting that the effectiveness of such programs may not solely depend on their interactivity but also on the quality and relevance of the information provided. Further research is needed to explore optimization strategies for such interventions, especially for persons with low PA, including user engagement, behavior change techniques, and the integration of objective PA tracking methods.

* Corresponding author.
E-mail address: Urs.Fichtner@uniklinik-freiburg.de (U.A. Fichtner).
[1] Shared first authorship.
[2] Shared last authorship.

https://doi.org/10.1016/j.invent.2024.100762
Received 24 June 2024; Received in revised form 30 July 2024; Accepted 31 July 2024

Trial registration: German Clinical Trials Register DRKS00020249; https://drks.de/search/en/trial/DRKS00020249.

1. Introduction

Physical inactivity is a major driver of the main causes of death worldwide. It is associated with increased risk of arterial hypertension, type 2 diabetes, premature cardiovascular diseases and mortality, higher risk of cancer, musculoskeletal disease, mental and cognitive health impairment and all-cause mortality (World Health Organization, 2020). The WHO and other medical associations recommend engaging in at least 150 min of moderate-intensity aerobic physical activity (PA) per week (equivalent to 30 min a day on most days) for health prevention. Additional muscle strengthening training on two days a week is associated with further health benefit (World Health Organization, 2020; Pelliccia et al., 2021). PA shows positive effects on the regulation of body weight, blood pressure, blood glucose and cholesterol levels, and thus reduces the incidence of diabetes, cardiovascular diseases and overall mortality. Results from meta-analyses, systematic reviews and large longitudinal studies indicate that PA reduces stress hormones, promotes healthy sleep patterns, and positively affects symptoms of mental illness (Pelliccia et al., 2021; Wen and Wang, 2017; Jurik and Stastny, 2019; Schuch et al., 2017; Ahmadi et al., 2022; Chen et al., 2022; Blodgett et al., 2023; Posadzki et al., 2020). According to the WHO, increasing PA could prevent four or five million deaths worldwide every year (World Health Organization, 2020). In 2015, about 55 % of adults living in Germany reported engaging in <150 min of moderate PA per week and fewer than 25 % followed the recommended combination of endurance and strengthening exercises (Finger et al., 2017). In nations of high sociodemographic index, sedentary lifestyle and associated disability-adjusted life years (DALYs) were growing between 2010 and 2019 (GBD 2019 Risk Factors Collaborators, 2020). A lower level of education and lower socio-economic status are correlated with less time spent on PA, overall poorer health behavior, and higher morbidity and mortality. In Germany, this also manifests itself regionally (Finger et al., 2017).

Web-based PA programs could effectively motivate a broad part of the population to engage in PA, including those with limited time or financial resources or who reside in areas with limited access to sports facilities. According to a cross sectional study of $n = 1014$ internet users in Germany in 2020, 57 % of the participants stated to use digital media for health related subjects and about 22 % ($n = 220$) used digital media in the context of their PA (De Santis et al., 2021).

Controversial results can be found in the literature. Some studies report promising results about effects of web-based PA interventions, other findings are rather disillusioning. Two meta-analyses examined the effects of PA-interventions. Jahangiry et al. included 15 studies and found a small increase in walking time due to web-based PA interventions compared to the control groups amounting a weighted mean difference of 0.17 min/week (CI [0.08; 0.27], effect size z = 3.47; p = .001) (Jahangiry et al., 2017). Effects of single studies within this meta-analysis are reported between −20 min/week (CI [−35.80; −4.20]) to +172 min/week (CI [79.70; 264.30]) (Jahangiry et al., 2017). Kwan et al. (2020) focused their meta-analysis on studies including people with a mean age > 50 years. The mean difference in PA time of nine included studies amounted to +53.20 min/week (CI [30.18; 76.21], effect size z = 4.53; $p < .00001$) (Kwan et al., 2020). The results of the single studies included here range between −300 min/week CI [−973.83; 372.23] and 204.5 min/week CI [−150.89; 559.89]. Significant single results are reported with mean differences in self-reported general PA ranging from +38.8 min/week (CI [1.80; 75.80]) and + 166.60 min/week (CI [1.55; 331.65]) (Kwan et al., 2020).

Both meta-analyses show that the wide variation in intervention effectiveness is linked to differences in the content of web-based interventions, study designs (including the timing of follow-up data collection), and participant numbers. Earlier interventions provided automated, standardized instructions and/or automated telecounseling. Later interventions increasingly applied individualized digital PA programs, offering tailored feedback based on users' recorded data. This data was partly tracked automatically via wearables and partly entered manually by users. The content of PA online-programs often follows principles of behavior change techniques and the most recent programs combine a variation of intervention modules. Written information, video exercise instructions, supportive motivational videos, electronic reminder, monitoring instruments, tailored feedback etc. are provided (Jahangiry et al., 2017; Kwan et al., 2020). Both meta-analyses point out that self-monitoring, supported by pedometers or wearables, might enhance PA time. Studies confirm that the more individualized and interactive online interventions are, the stronger their effects are on PA (Kwan et al., 2020; Oosterveen et al., 2017; Foster et al., 2013).

Technical progress may facilitate the interactivity and individual tailoring of web-based PA programs. These attributes might enhance user's commitment to the program, as some studies indicate (Mclaughlin et al., 2021). Nevertheless, McLaughlin et al. show a weak overall relationship between engagement in interactive digital health interventions and PA outcomes (estimated standardized regression coefficient of eleven included studies = 0.077 (CI [0.153; 0.347] effect size z = 2.36; p = .002) (Mclaughlin et al., 2021). They also emphasize the high heterogeneity of studies included. While one of these reports a very small, negative effect of engagement on PA outcome, the strongest dose-response relationship amounts to 0.250 (CI [0.153; 0.347]). The latter intervention, investigated by Davies et al., combined various methods to improve self-monitoring and motivation, like pedometers, team and individual challenges, and virtual walking buddies (Davies et al., 2012). This corresponds to other studies, reporting higher effects on PA by the offer of self-monitoring options (Kwan et al., 2020; Oosterveen et al., 2017; Foster et al., 2013; Mclaughlin et al., 2021).

This overview shows that web-based PA programs have the potential to enhance PA. However, their effectiveness was found to be very heterogeneous and related to various aspects. Technical progress and digitalization offer possibilities to improve tailored feedback and to increase the usability of complex programs.

This article presents the evaluation of a 12-week multimodal web-based PA program, as a part of a comprehensive TK-HealthCoach, offered by the German health insurance company Techniker Krankenkasse (TK). The program was developed by an expert team comprising physicians, psychologists, exercise and nutrition scientists, addiction specialists, and IT experts. The TK-HealthCoach has an interactive, web-based format and consists of three individual programs that separately pursue the three health goals of (a) increasing fitness, (b) losing and maintaining weight, and (c) smoking cessation. Accordingly, three concurrent evaluation studies were conducted by an interdisciplinary research team at the University of Freiburg, Germany (see study protocol (Tinsel et al., 2021). Effects of the smoking cessation coach and the weight loss coach are already published (Fichtner et al., 2022; Maiwald et al., 2023; Kohl et al., 2022; Brame et al., 2023; Kohl et al., 2023). This study examines the fitness-related program (TK-FitnessCoach), which aims to increase the PA of the users. We implemented a randomized controlled trial and compared the TK-FitnessCoach, as an interactive web-based health program (intervention group, IG), with non-interactive web-based health information (control group, CG) to evaluate the effectiveness of both programs.

The Primary outcome is the effect on self-reported PA. Secondary endpoints are the effects on goal attainment, health-related quality of life, weight, and eating behavior at post intervention, 6 and 12 months after the end of the 12-week program. We hypothesized (1) statistically significant improvements of PA one year after the progam with small to medium effect sizes in both groups and (2) stronger effects on primary and secondary outcomes in the intervention group (TK-FitnessCoach) in comparison with the control group with small to medium effect sizes.

We also investigated whether the intensity of use of the TK-FitnessCoach has an impact on its effects (dose-response relationship). Our third hypothesis (3) was that study participants in the intervention group using the program more intensively achieve stronger outcome improvements than those who accessed the program less intensively.

Furthermore, we asked the participants if they used other online fitness programs in addition to the TK-FitnessCoach.

2. Methods

2.1. Study design

This study was part of an evaluation project addressing a German-speaking adult population interested in improving their health behavior. Within five online two-arm randomized controlled trials (RCTs) of parallel design, we evaluated the effectiveness of three health coaches with three health goals: weight loss and maintaining weight (Fichtner et al., 2022), smoking cessation (Maiwald et al., 2023) and increasing fitness. Two sub studies examined the impact of the weight loss program (Kohl et al., 2022; Kohl et al., 2023) and the fitness program (Brame et al., 2023) on medical parameters.

In this study, we evaluated the effectiveness of the fitness program based on pre-defined health outcomes. Participants allocated to the intervention group (IG) received a web-based, interactive, tailored intervention (TK-FitnessCoach), whereas the control group (CG) had access to a static homepage with evidence-based information on PA and training. IG participants were recommended to use the TK-FitnessCoach for 12 weeks.

This RCT encompasses four measurement time points: T0, representing the baseline assessment conducted prior to randomization; T1, which was the post-intervention evaluation, conducted at 12 weeks subsequent to the commencement of the program; T2, denoting the follow-up appraisal carried out six months after the program's conclusion; and T3, indicating the final assessment conducted 12 months after the conclusion of the program.

Randomization was executed via permuting block randomization with variable block sizes of 4, 6 and 8. For this purpose, a randomization list was created via the RITA software (Pahlke et al., 2004). Participants were automatically randomized after completing the baseline questionnaire (T0). Participants got invitation e-mails to answer the questionnaires at every time point. In case of no response, they were reminded via e-mails after 13 days, four weeks and in week 7 before getting blocked after 8 weeks. This procedure was applied to all time points.

The study protocol was approved by the Ethics Committee of Albert-Ludwigs-University, Medical Center, Freiburg (vote no. 237/19). This study is registered in the German Clinical Trials Register https://www.drks.de/drks_web/ with the registry ID DRKS00020249 at 11 December 2019. Recruitment started at 1 January 2020. Informed consent was obtained from all individuals participating in this study.

2.1.1. Intervention

Participants of the intervention group received a web-based interactive, tailored, and technically guided intervention for a period of 12 weeks aiming at increasing PA. The TK-FitnessCoach is theoretically based on the MoVo concept, which implements motivational and volitional strategies to change health behavior (Fuchs et al., 2009) and is structured in three consecutive phases: 1. Getting to know and trying out (week 1–3); 2. Establishing new behavior (week 4–6); 3. Strengthening and perpetuating habits (week 7–12). This online coach is structured in modules and at the first session, all users are guided through a comprehensive algorithm-based test to check if they are eligible to set the basis for their individually tailored coaching content. For finding a suitable PA as well as the right level of intensity, users can participate in a fitness test at the beginning. In a next step, participants have the option to individualize their coaching by selecting and planning physical activities they would like to try in the following weeks. 16 different activities including plans with various intensity levels for 12 weeks are available (e.g. strength workouts, flexibility training, endurance plans for running, walking, and cycling, Pilates workouts, yoga). In addition, text-based lessons are provided to further deepen the knowledge about training and PA. Users can further supplement their individual fitness program with modules to improve their diet in order to receive additional support. Participants receive up to three follow-up prompts to continue with the program after one week of inactivity. The users can deactivate these.

To conduct this RCT, a quasi-intervention for the control group was designed. Participants of the control group received access to a static web page with evidence-based information about PA and training, similar to the text-based lessons of the TK-FitnessCoach. There were no further recommendations or training plans provided for the control group. The TK-FitnessCoach and the static web page of the control group were developed in a responsive web design to ensure the best user experience for both desktop and mobile users. More information on the coach can be found elsewhere (Tinsel et al., 2021; Brame et al., 2022).

2.2. Participants and recruitment

To recruit participants, several campaigns were launched across various online and print media from January 1st 2020 to September 28th 2020. Deadline for the last inclusion was October 5th 2020 and end of data collection was January 10th 2022. Participants were rewarded with a €25 Amazon coupon code. Due to high dropout rates after the first measurement point (T1), the incentivization strategy was revised during the study to offer €10 coupons for each measurement time point instead of the €25 at the end of the study.

Healthy persons of any gender aged 18 years or older were included in this trial regardless of their health insurance. Exclusion criteria were: taking part in another study aiming to change behavior towards fitness, health impairments that make the program not suitable, pregnancy and being currently underweight or obese (BMI > 40, waist circumference > 200 cm). Patients were informed that there must be no health impairments, which, according to medical assessment, partially or completely preclude participation in the health coaching and its modules. It was possible to participate in only one of the three online programs provided for evaluation at the same time. Sample size calculations were made in G*Power (Faul et al., 2007) for small effect sizes ($d = 0.20$) with $1\text{-}\beta = 0.80$ and $\alpha = 0.50$. We assumed a dropout rate of 50 %. Based on our sample size calculation, we aimed for $n = 1114$ participants at baseline. Detailed information can be found in the study protocol (Tinsel et al., 2021).

2.3. Primary outcome

At each measurement time point, an online questionnaire was used to measure variables that were considered relevant for statistical modeling and prediction of effectiveness.

Physical activities were assessed with the Physical Activity, Exercise, and Sport Questionnaire (Bewegungs- und Sportaktivität Fragebogen, BSA-F, (Fuchs et al., 2015). We used the *Physical Activity* index as primary outcome for this study, which is the sum of walking, cycling, and sport activities displayed by the *Sport Activity* index (see below). We did not use other components of the BSA-F such as climbing stairs, care activities, or cleaning activities, because these components were not addressed in the TK-FitnessCoach. The *Sport Activity* index was calculated based on three open-ended questions. Participants could enter a maximum of three exercises or sport activities they had regularly performed within the last 4 weeks, and were then asked to indicate the frequency and duration (in minutes) of each activity episode. Not all activity entries were considered sport activities (e.g., walking the dog, gardening, sexual activities, or job-related activities). The data processing of text entries was conducted using an application developed specifically for this purpose (Sehlbrede and Wurst, 2022). This application features a dictionary that classifies sport activities into different categories, such as winter sports, water sports or non-sporting activities. All entries that were not categorized as sport activity were excluded from the calculation of the sport activity index. The *Sport Activity* index was created as a sum of the products of frequency and duration per activity. For each sport activity, the number of minutes per day was multiplied by the frequency. The number of minutes per month was divided by 4 to calculate the index representing the PA in minutes per week. If there was more than one sport activity entry, this was computed for each entry and then all index values were summed up. The maximum time per single activity was truncated to 120 min per episode because longer lasting sport activities such as hiking or downhill skiing typically contain longer resting periods. The maximum number of days allowed was 30 times per month.

2.4. Secondary outcomes

Body weight of the participants was assessed using self-reported data in kilograms. *Goal Attainment* (self-developed) was measured with a single six-point Likert item with the categories "worse than before", "no change", "25 % achieved", "50 % achieved", "75 % achieved", "100 % achieved", "$\geq$100 % achieved". We coded the categories with values from -1 to 5, while no change was coded with 0. *General Health* was assessed with one single item where participants should rate their health in comparison to another person of the same age and gender on a scale from 0 ("substantially worse") to 10 ("substantially better") (Renner et al., 1996). Eating behavior was measured with the German eating behavior scale, which has two subscales: *health-conscious eating behavior* (*hc-EB*) and *weight-controlling eating behavior* (*wc-EB*) (Wurst et al., 2022). We calculated a mean score for each subscale, which can range between 1 and 5 with higher values indicating a more conscious or controlling eating behavior, respectively. Reliability was satisfactory and ranged between ω_t=0.86 and 0.89 for both subscales. Health-related *Quality of Life* (QoL) was assessed with the German version of the SF-12, which is a short form of the SF-36 (Wirtz et al., 2018). We calculated a score for physical (SF12-K) and mental quality of life (SF12-P), both scores ranged between 0 and 100 with higher values indicating higher quality of life.

2.5. Confounders

With the rehabilitation-specific comorbidity score, we assessed the amount and severity of *Health Impairments (*Glattacker et al., *2007)*. The score ranges between 0 and 10 with higher values indicating more frequent and more severe comorbidities. Reliability was ω_t=0.51 which is not satisfactory. *Mental Health* was assessed with two items of the German version of the PHQ-D (Löwe et al., n.d.). We aggregated both items to a mean score ranging between 1 and 4, with higher values indicating lower mental health. Sport and Exercise related *Social Support* was measured with the scale Social Support consisting of 12 items using a 4-point Likert response scale (Fuchs, 1997). We calculated a mean score over all items with higher values indicating higher social support. Reliability was satisfactory for all time points with ω_t ranging between 0.86 and 0.88.

We applied a latent class analysis to categorize all users based on their weekly login behavior and found five user types: constant user, halftime user, partial user, rare user and one-time user. This categorization was used to investigate the dose-response association in the intervention group as stated in hypothesis 3. User types differ in their login behavior per week: constant users showed higher login frequency over all 12 weeks, halftime users showed logins up to week seven, partial users had sporadic logins over all 12 weeks but also weeks with no logins, rare users had logins only at the first two weeks and one-time users had only one login at all. Hypothesis 3 was only tested with the intervention group.

2.6. Samples

We used different kind of samples in our analysis. Our primary sample is the intention-to-treat (ITT) sample. This sample contains all cases ($n_{IG} = 1,153$, $n_{CG} = 1,177$), that were randomized and had plausible values. We used multiple imputation to impute missing values in this sample (see section missing data and imputation). Analyses using this imputed ITT sample served as primary method for testing hypotheses 1–3. Our second sample covers complete cases (CC) only. This sample ($n_{IG} = 258$, $n_{CG} = 363$) fulfills two different purposes: first it represents our analysis per protocol as defined a priori in the study protocol (Tinsel et al., 2021). For this sample, information was available at all measurement time points. Second, the CC analysis serves as sensitivity analysis for the imputation as advised versa (van Buuren, 2018). Furthermore, we used the raw ITT sample without imputation for additional sensitivity analyses. We limited this sample to those, who had more than one login into the coach (defined as multi-time-users, MTU). The sample size for MTU was $n = 741$ in the intervention and $n = 458$ in the control group and we applied no additional imputation procedure for the MTU sample.

2.7. Missing data and imputation

Incomplete data were imputed under full conditional specification using the *mice* package in R (van Buuren, 2018). Specifically, we used the two-stage method for a normal model with heteroscedastic errors (Resche-Rigon and White, 2018), which is implemented in the R package *micemd (*Audigier and Resche-Rigon, 2021*)*. We imputed 100 data sets and pooled the results over all data sets with the package *mitml (*Grund et al., *2019)*. We calculated the relative increase in variance (RIV) and fraction of missing information (FMI) for the imputed models as measures for the influence of the missing data on the estimated model parameters (Enders, 2010) and performed sensitivity analysis.

2.8. Statistical procedure

All statistical analyses were conducted with the free statistical software *R*, Version 4.0.5. Visualizations were created with the R package *ggplot 2* (Wickham, 2016). To test our hypotheses, we used multilevel generalized linear mixed models (GLMM) with person ID as cluster variable and the count of minutes of PA as outcome variable. To predict the secondary outcomes, we applied a linear mixed model. The predictors on level 1 were change over time (T0 to T3), group mean centered mental illness and social support. On level 2, we used grand mean centered age, general health, health impairments as well as the categorical predictors gender and group differences between intervention control group. Hypothesis 3 was tested only with the intervention group; therefore, we dropped the study group predictor variable.

We checked for overdispersion and zero inflation with R package *DHARMa (*Hartig, *2022)* and because both tests were positive, we changed the model to a zero-inflated negative binominal mixed model, using R package *glmmTMB* (Brooks et al., 2017). To calculate linear mixed models we used the R package *lme4* (Bates et al., 2015) together with the package *lmerTest* (Kuznetsova et al., 2017) to determine Satterthwaite degrees of freedom. As effect size we calculated marginal and conditional R^2 based on (Nakagawa et al., 2017). Pairwise comparisons with Tukey-adjustment for multiple testing implemented in the package *emmeans* (Lenth, 2021) were used to compare different time points. Differences between intervention and control group in goal attainment were assessed with Welch's *t*-test. Goal attainment was not measured at baseline, so we focused on differences in goal attainment at the time points post intervention (T1) and 12-month follow up (T3). As effect size for mean differences between and within groups we used Cohen's *d* with confidence intervals calculated with the package *effectsize (*Ben-Shachar et al., *2020)* and used the interpretation of $| > 0.30|$ as small, $| > 0.50|$ as medium and $| > 0.80|$ as large effect.

3. Results

3.1. Missing data patterns

We assumed that the structure of missing values will follow the MAR (missing at random) assumption and investigated the pattern of missing data following guides from (van Buuren, 2018; McKnight, 2007). We identified two different kinds of missing values: missing values due to complete dropout (=completely missing data) for one or more time points after baseline and missing values on single items for a time point without completely missing. Table 1 shows the patterns of completely missing data for each time point. The most frequent patterns were "T1, T2 & T3 completely missing" (48.6 %, 38.0 %) and "T2 & T3 completely missing" (7.6 %, 12.2 %). Patterns of completely missing for one time point or two non-consecutive time points were not very frequent ($\leq$ 7 %). In addition, not all participants, who filled out the survey at all time points had complete data: 90 % (n = 258/1153) of the cases in the intervention group (IG) and 94 % (n = 363/1177) of the cases in the control group (CG) had no missing data on all relevant variables if they took part at every time point.

Some of the completely missing data after baseline could be explained by information from a short active deregistration survey after participants manually signed off from the study. For 481 cases we had information about the reason for quitting, like personal reasons (IG = 20 %, CG = 11 %), program specific reasons (IG = 65 %, CG = 75 %) or technical problems (IG = 15 %, CG = 14 %). The most frequently cited reasons were that the program did not provide new inspiration (20 %), did not meet expectations (21 %), and was too complicated (14 %).

Baseline could not be completely missing because taking part at baseline was mandatory to be randomized in the study. Survey timestamp was a variable, which coded automatically the actual date when the online survey was accessed. If cases did not have a date entry at one of the time points, they were considered as completely missing.

3.2. Sample size

In total, 3108 persons registered for the TK-FitnessCoach study and 2584 (83 %) responded to the final invitation mail and were randomized to either the IG or CG (Fig. 1). For this evaluation, we excluded those persons who participated in an additional medical sub study. This sub group received additional medical examinations and activity trackers which might have had biased our trial (Brame et al., 2023). We also excluded all cases with implausible or invalid values. These were defined as changes in age > 1 or decrease in age, gender changes over measurement points, or zero variance on the two subscales of eating behavior (hc-EB, wc-EB), on the variables health impairments, both quality of life subscales (SF12-K, SF12-P), social support and mental health, because zero variance indicates an invalid response pattern.

After deletion of implausible data (3 % of the randomized cases), there were 1153 cases left in the IG and 1177 cases in the CG for ITT analyses.

Table 1
Patterns of missing data for each time point based on survey timestamp.

	Intervention group	Control group
Information for all time points available	286 (24.8 %)	386 (32.8 %)
Only T1 completely missing	62 (5.4 %)	44 (3.7 %)
Only T2 completely missing	22 (1.9 %)	29 (2.5 %)
Only T3 completely missing	48 (4.2 %)	78 (6.6 %)
T1 & T2 completely missing	42 (3.6 %)	24 (2.0 %)
T1 & T3 completely missing	46 (4.0 %)	25 (2.1 %)
T2 & T3 completely missing	87 (7.6 %)	144 (12.2 %)
T1, T2 & T3 missing	560 (48.6 %)	447 (38.0 %)

Note. T1 = post-intervention, T2 = 6-month follow-up, T3 = 12-month follow-up.

3.3. Sample description

We achieved equally distributed sociodemographic values in both the IG and the CG indicating a well-implemented randomization. Table 2 shows that the majority of the participants were female (75 %) and the mean age over all participants was 43 years. In total, 72 % of the sample in the IG (respectively 68 % in the CG) were married or living in a steady partnership. Three quarters of the respondents were employed (IG: 77 %, CG: 76 %). Most of the participants (70 % in the IG and 67 % in the CG) were insured with the *Techniker* health insurance (TK) and had no experience with online health programs (93 % in the IG and the CG). Approximately one fifth of the sample (22 % in CG/IG) reported an earlier or actual diagnosed mental disorder. About 15 % of the respondents in each group considered themselves as smokers and tried to increase their fitness in the past.

From descriptive perspective, the primary outcome (PA) was lowest on average in both groups for T0 with 278 min/week in the IG and 273 min/week in the CG (see Fig. 2, Table 3). The highest mean PA was measured in both groups at T1 (IG: 357 min/week; CG 356 min/week). While we observed a small decrease at T2 (348 min/week) and a small increase at T3 (350 min/week) in the IG, a decrease was found in the CG, which remained stable over the two time points T2 and T3 (347 min/week). A similar pattern was observed for the different components of PA: The lowest mean activity in min/week was found in both groups for T0 with an increase at T1 and slight decrease at T2/T3. Sports activity increased in the mean at T3 for both the IG (116 min/week) and the CG (115 min/week). Cycling activity increased in the mean at T3 in the CG (67 min/week).

From a descriptive perspective, secondary outcomes improved slightly on average. Eating behavior enhanced little and remained at a higher level. Weight loss amounted to about 1 kg on average between T0 and T1, and participants maintained a lower level of weight on average until T3. Participants reported from T1 to T3 that they attained their goals in a mean of about 50 % (corresponds to the value 2). Quality of life improved scarcely on average. No notable differences between the two groups or changes over time can be reported. This also applies to the covariates (see Table 3).

3.4. Uni- and bivariate analyses

The mean difference for the long-term effect (T0-T3) on PA translated into an effect size of $d = 0.33$ (99 %-CI [0.24, 0.46]) for the IG, and $d = 0.34$ (99 %-CI [0.23, 0.44]) for the CG (Table 4). Short-term effects (T0-T1) in the IG were slightly higher ($d = 0.35$, 99 %-CI [0.24, 0.45]) than the long-term effects. All time-related effects for PA can be considered small effects. Between group differences were extremely small at T1 ($d = 0.004$, 99 %-CI [−0.08, 0.09] and at T3 ($d = 0.02$, 99 %-CI [−0.06, 0.10]). These findings do not differ substantially between ITT, CC and multi-time-user subsamples.

For all secondary outcomes, the effect sizes were smaller than the cut-off of 0.3 for Cohen's *d*, which indicate no long-term and no short-term effects for both the IG and the CG on the secondary outcomes. Goal attainment, which was tested with the Welch-Test due to its scale level, showed also no significant differences between the IG and the CG at T1 ($t_w(2243.8) = 0.39$, $p = .698$, d = 0.02 95 %-CI [−0.07, 0.10]). This was also the case at T3 ($t_w(2286.1) = 0.30$, $p = .768$, d = 0.01 95

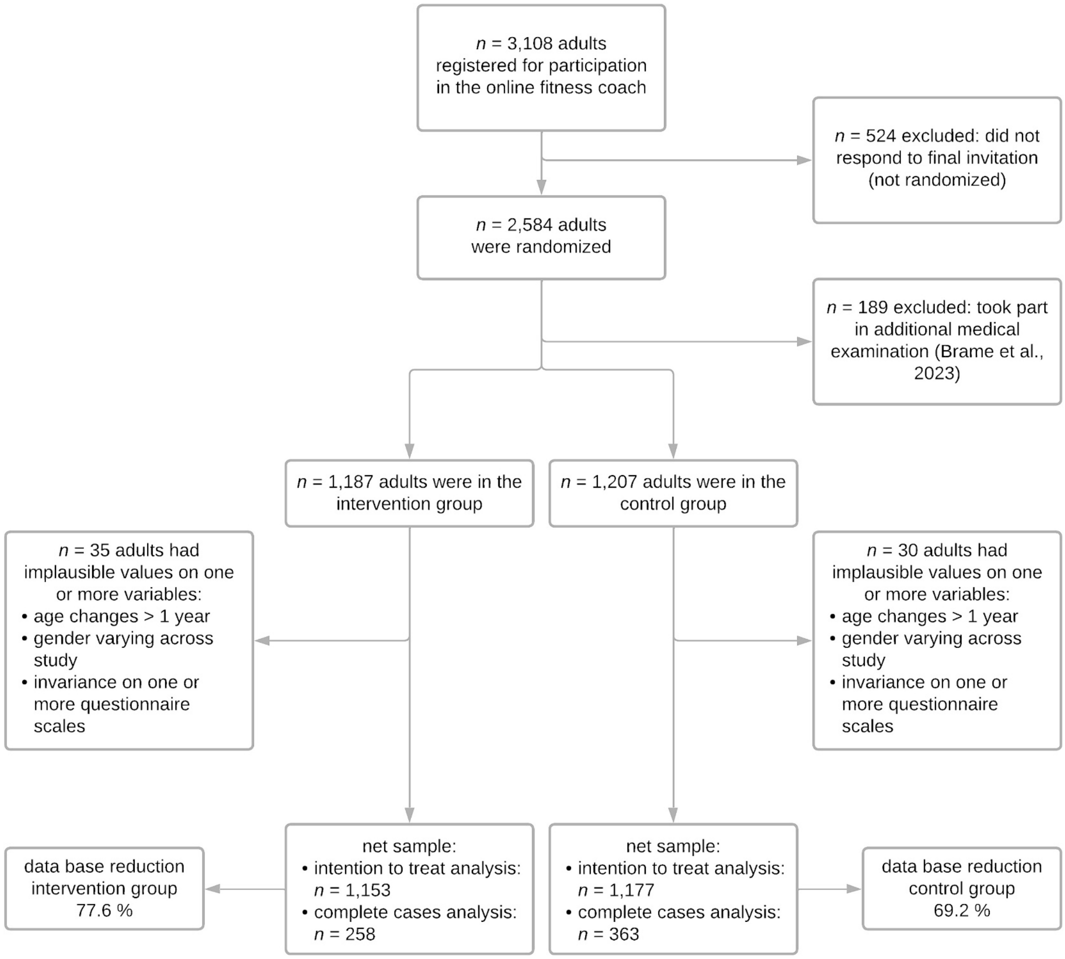

Fig. 1. Participant flow for the TK-FitnessCoach evaluation study.

Table 2
Sample characteristics (ITT sample).

	Intervention group	Control group
	(n = 1153)	(n = 1171)
	T0	T0
	Sample statistics	
Age (*M*, [*SD*])	42.80 [13.94]	43.14 [13.90]
Gender (%)		
Female	76.1	74.7
Male	23.9	24.8
Diverse	–	0.5
Family status (%)		
Unmarried / divorced / widowed / living separately	28.0	21.7
Married / in steady partnership	72.0	68.3
Job status: employed (%)	76.5	76.4
Prior experience with online health-coaches (yes, %)	6.8	6.8
Earlier/ actual mental disease diagnosis (yes, %)	21.6	22.2
Medical insurance (%)		
TK	69.9	66.9
Private insurance	8.8	9.4
Other	21.2	23.5
Smoker (yes, %)	14.5	14.7
Tried to increase fitness in past (no, %)	14.3	15.3

Note. M = arithmetic mean, *SD* = standard deviation. Empty cells: was not measured at this time point.

%-CI [−0.07, 0.09]).

3.5. *Multilevel analyses*

The results from the linear mixed models are displayed in Table 5. The three different models predict the primary outcome PA using the pre-specified predictors and confounders for the imputed ITT data set, the CC data set and for the reduced multi-user data set. In all three models, we found a significant increase in PA over time for both groups. However, we found no significant difference between IG and CG, expressed by the non-significant interaction between group and time as well as the non-significant coefficient for group membership. Furthermore, general health as well as social support showed a beneficial impact as confounder for both groups in all three data sets. Age, gender and health impairments had no significant effect on PA at T3.

For the secondary outcomes (see Table 6), we also did not observe any effects that were explained by group differences between IG and CG. Although healthy eating (hc-EB and wc-EB) increased over time, weight was not changing significantly. We found following associations: Increase in healthy eating (hc-EB and wc-EB) was associated with older age, male gender and general health. Higher weight at T3 was associated with older age, male gender, and lower health status. A significant decrease in weight-controlling eating behavior (wc-EB) was found for lower mental health. Age was associated with QoL outcomes, with older age predicting lower physical health-related QoL and higher mental health-related QoL. Men showed lower mental health-related QoL than women. Higher general health and mental health were positively associated with health-related QoL. Social support showed, in addition to its

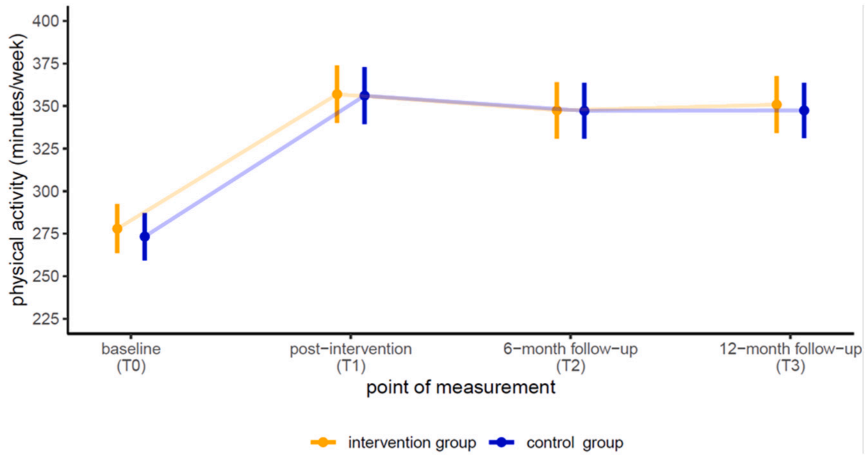

Fig. 2. Mean changes on PA with 95 %-confidence intervals (ITT sample with $n_{IG} = 1,153, n_{CG} = 1,177$).

Table 3
Distribution of primary and secondary outcomes and confounders (ITT sample).

	Intervention group (n = 1153)				Control group (n = 1171)			
	T0	T1	T2	T3	T0	T1	T2	T3
Primary outcome								
Physical Activity (min/week; *M*, [*SD*])	277.92 [248.76]	356.97 [292.55]	347.49 [285.62]	350.84 [289.44]	273.30 [242.00]	356.10 [290.93]	347.19 [286.63]	347.42 [284.22]
Physical Activity (min/week; *Md*, [*SD*])	210.00 [192.74]	328.39 [116.77]	322.30 [110.98]	327.02 [105.97]	210.00 [200.15]	322.50 [137.14]	323.20 [127.87]	316.18 [115.90]
Components of physical activity								
Sports Activity (min/week; *M*, [*SD*])	83.12 [124.86]	114.77 [142.85]	110.93 [143.81]	115.47 [155.06]	84.10 [131.23]	119.54 [152.54]	107.05 [141.07]	114.93 [149.14]
Movement Activity (Walking and cycling) (min/week; *M*, [*SD*])	194.80 [194.72]	240.91 [221.51]	235.49 [212.09]	235.53 [213.34]	189.20 [176.25]	236.17 [208.20]	237.98 [210.49]	235.49 [209.07]
Only walking (min/week; *M*, [*SD*])	142.03 [166.35]	171.21 [176.23]	175.03 [176.03]	171.42 [177.35]	133.13 [143.31]	165.66 [174.86]	175.45 [176.89]	167.12 [169.62]
Only cycling (min/week; *M*, [*SD*])	52.77 [90.14]	67.32 [110.84]	56.63 [101.37]	65.13 [106.96]	56.07 [93.89]	70.84 [114.13]	61.20 [103.62]	67.17 [122.56]
Secondary outcomes								
Goal Attainment (*M*, [*SD*])		2.12 [1.50]	2.06 [1.55]	2.08 [1.50]		2.10 [1.55]	2.04 [1.56]	2.07 [1.56]
Body weight (kg; *M*, [*SD*])	76.79 [16.86]	75.80 [16.15]	75.76 [16.03]	75.67 [16.02]	77.02 [16.54]	76.14 [15.86]	76.06 [15.84]	76.10 [15.91]
SEV hc-EB (*M*, [*SD*])	3.38 [0.62]	3.49 [0.61]	3.49 [0.61]	3.50 [0.61]	3.37 [0.65]	3.50 [0.61]	3.52 [0.61]	3.51 [0.61]
SEV wc-EB (*M*, [*SD*])	2.56 [0.73]	2.66 [0.74]	2.66 [0.72]	2.64 [0.72]	2.54 [0.71]	2.69 [0.73]	2.68 [0.73]	2.66 [0.74]
SF12-K (*M*, [*SD*])	48.83 [7.75]	49.10 [8.35]	48.78 [8.74]	48.86 [8.58]	48.74 [7.92]	49.05 [8.49]	48.84 [8.62]	48.99 [8.57]
SF12-P (*M*, [*SD*])	42.20 [10.29]	43.00 [10.54]	42.54 [10.82]	42.96 [10.85]	42.11 [10.74]	43.01 [10.70]	42.35 [10.92]	42.54 [10.90]
Confounders / Covariates								
General Health (*M*, [*SD*])	5.81 [2.12]				5.90 [2.10]			
Health Impairments (*M*, [*SD*])	0.23 [0.30]				0.23 [0.29]			
Mental Health (*M*, [*SD*])	1.70 [0.63]	1.70 [0.63]	1.71 [0.64]	1.71 [0.65]	1.69 [0.63]	1.70 [0.65]	1.72 [0.65]	1.71 [0.66
Social Support (*M*, [*SD*])	2.63 [0.48]	2.63 [0.50]	2.63 [0.48]	2.63 [0.49]	2.64 [0.49]	2.63 [0.50]	2.64 [0.50]	2.62 [0.49]

Note. M = arithmetic mean, *SD* = standard deviation, Md = Median. Empty cells: was not measured at this time point. Physical activity is the sum of movement and sport activity. Movement activity is the sum of walking and cycling activity. Scale range: Eating behavior subscales hc-EB/wc-EB 1–5, Quality of Life (SF12 subscales) 0–100, General Health 0–10, Health Impairments 0–10, Mental Health 1–4, Social Support 1–4, Goal Attainment: −1 = worse than before, 0 = no change, 1 = 25 % achieved, 2 = 50 % achieved, 3 = 75 % achieved, 4 = 100 % achieved, 5 ≥ 100 % achieved.

association with PA, a positive correlation with mental health-related QoL. However, mental health related QoL was not associated with PA. The amount and severity of health impairments was negatively associated with health-related QoL.

3.6. Dose-response association in the IG

We found no significant interaction effect for the user type ($\chi^2(12) = 10.62$, $p = .562$) on PA and no significant main effect on PA for differences between user types ($\chi^2(4) = 3.65$, $p = .455$) in the IG. Pairwise comparison showed no significant differences between the user types regardless the time point. Short-term and long-term effects were significant for all user types except for the partial users. The mean difference for the partial users were comparable to the change for other groups, the missing significance might be a result of the rather small group size in comparison to all other groups. Fig. 3 shows change in PA based on marginal means from pairwise comparisons with 95 % confidence intervals.

Table 4
Cohen's d effect sizes with 99 %-confidence intervals for short- and long-term effects on primary and secondary outcomes (ITT sample).

Outcome	Short-term (T1-T0)		Long-Term (T3-T0)	
	Intervention group	Control group	Intervention group	Control group
Physical Activity[a] (min/week)	0.35 [0.24, 0.45]	0.35 [0.25, 0.46]	0.33 [0.22, 0.44]	0.34 [0.23, 0.44]
Weight[b] (kg)	−0.07 [−0.18, 0.04]	−0.06 [−0.17, 0.04]	−0.08 [−0.19, 0.03]	−0.07 [−0.17, 0.04]
Health conscious eating behavior[a]	0.20 [0.09, 0.31]	0.23 [0.13, 0.39]	0.22 [0.12, 0.33]	0.25 [0.14, 0.35]
Weight controlling eating behavior[a]	0.17 [0.06, 0.28]	0.24 [0.14, 0.35]	0.14 [0.03, 0.24]	0.20 [0.09, 0.31]
Physical health related QoL[a]	0.04 [−0.07, 0.15]	0.04 [−0.06, 0.15]	0.01 [−0.10, 0.11]	0.04 [−0.07, 0.07]
Mental health related QoL[a]	0.09 [−0.02, 0.20]	0.10 [−0.01, 0.20]	0.09 [−0.02, 0.20]	0.06 [−0.05, 016]

Note. We chose a 99 % confidence interval [square brackets], because there are a total of six possible contrasts/differences between time points, so we calculate 1–0.05/6 to adjust the confidence interval. T0 = baseline, T1 = post-intervention, T3 = 12-month follow-up. Interpretation: > 0.20 small effect, > 0.50 medium effect, > 0.80 large effect. Secondary outcome Goal Attainment is not shown because it was not measured at baseline (T0). QoL = Quality of life.
[a] Mean value increases.
[b] Mean value decreases.

Table 5
Results from mixed model analysis for Primary Outcome over time T0 to T3.

	Primary Outcome: physical activity[a]		
	ITT	CC	MTU
Fixed effects			
Intercept	**5.68 (0.03)**	**5.61 (0.07)**	**5.54 (0.05)**
Group [control]	−0.02 (0.04)	−0.05 (0.06)	−0.09 (0.05)
Time	**0.07 (0.01)**	**0.05 (0.02)**	**0.10 (0.02)**
Group * Time	0.01 (0.02)	0.02 (0.02)	0.02 (0.02)
Age[c]	−0.00 (0.01)	0.02 (0.03)	0.01 (0.02)
Gender [female]	−0.01 (0.03)	0.04 (0.06)	−0.00 (0.03)
General health[c]	**0.10 (0.01)**	**0.18 (0.03)**	**0.23 (0.03)**
Mental health[b]	−0.03 (0.03)	−0.06 (0.03)	−0.05 (0.03)
Social Support[b]	**0.12 (0.05)**	**0.22 (0.05)**	**0.19 (0.05)**
Health impairments[c]	0.01 (0.02)	−0.01 (0.03)	0.04 (0.02)
Random effects			
Residual	–	–	–
Intercept (σ^2)	0.099	0.308	0.323

Note. Sample for primary outcome: ITT = intention to treat, CC = only complete cases, MTU = only multi-time-user. Regression coefficients with standard errors in parenthesis. Reference group for categorical predictors in square brackets.ITT data was imputed and pooled over all 100 imputed data sets. Bold: $p < .05$.
[a] Zero-inflated negative binominal mixed model.
[b] Group mean centered.
[c] Grand-mean centered.

3.7. *Usage of other PA programs*

Our exploratory analyses showed that some participants used other online health programs or tools at the beginning of the study (see Fig. 4). Though the share is low, there were 6.7 % persons in our sample using another online fitness program.

Since the amount of study participants who used other online health programs was highest for the category "online fitness programs", we further investigated the usage behavior during the whole study period for this offer (Fig. 5). In fact, the proportion of using other online fitness programs arose at T1 to 24.4 %, at T2 to 27.8 % and it decreased again at T3 to 20 %. However, the shares between the CG and the IG seemed to be quite similar. Due to too low cell count frequency, we could not statistically test for significant effects, but it seems that there is no real difference between CG and IG concerning the use of other online fitness programs. Therefore, the lack of significant differences between IG and CG in the primary outcomes is probably not attributable to a different use of third party programs in the two groups.

Table 6
Results from mixed model analysis for Secondary Outcomes over time T0 to T3.

	Secondary outcomes[a]				
	Body weight	hc-EB	wc-EB	SF12-K	SF12-P
Fixed effects					
Intercept	**82.79 (0.67)**	**3.27 (0.04)**	**2.52 (0.03)**	**49.31 (0.28)**	**43.24 (0.39)**
Group [control]	0.27 (0.60)	−0.00 (0.02)	−0.01 (0.03)	−0.15 (0.29)	−0.14 (0.38)
Time	−0.34 (0.23)	**0.04 (0.01)**	**0.02 (0.01)**	−0.02 (0.11)	0.20 (0.15)
Group * Time	0.06 (0.31)	0.04 (0.01)	0.01 (0.01)	−0.08 (0.15)	−0.06 (0.19)
Age[c]	**0.84 (0.26)**	**0.04 (0.01)**	**0.08 (0.01)**	**−0.65 (0.12)**	**0.93 (0.16)**
Gender [female]	**−8.32 (0.61)**	**0.18 (0.02)**	**0.10 (0.03)**	−0.49 (0.25)	**−1.10 (0.36)**
General health[c]	**−3.17 (0.27)**	**0.08 (0.01)**	**0.05 (0.01)**	**1.35 (0.13)**	**1.18 (0.16)**
Mental health[b]	0.08 (0.49)	−0.03 (0.02)	**−0.06 (0.02)**	**−1.09 (0.31)**	**−4.68 (0.41)**
Social Support[b]	−0.11 (0.88)	0.05 (0.04)	0.03 (0.04)	0.30 (0.50)	**1.63 (0.57)**
Health impairments[c]	0.63 (0.88)	−0.00 (0.01)	0.01 (0.01)	**−1.89 (0.13)**	**−1.89 (0.17)**
Random effects					
Residual	175.64	0.27	0.39	53.98	86.57
Intercept (σ^2)	63.37	0.09	0.12	8.6	18.75

Note. Samples for secondary outcomes was always ITT sample. hc-EB = health-conscious eating behavior, wc-EB = weight-controlling eating behavior, SF12-K = physical health related quality of life, SF12-P = mental health related quality of life. Regression coefficients with standard errors in parenthesis. Reference group for categorical predictors in square brackets. ITT data was imputed and pooled over all 100 imputed data sets. Bold: $p < .05$.
[a] Linear mixed model.
[b] Group mean centered.
[c] Grand-mean centered.

4. Discussion

4.1. *Principle results*

The objective of this study was to examine the impact of a 12-week interactive web-based health program (TK-FitnessCoach) on physical activity and health in adults. The interactive fitness program (IG) was compared with a non-interactive website containing evidence-based information to enhance physical activity (CG). In the ITT data set, both, the intervention group (n = 1153) and the control group (n = 1177) exhibited significant increases in PA over time ($\Delta IG_{(T3\text{-}T0)}$ = 72.92 min/week; $\Delta CG_{(T3\text{-}T0)}$ = 74.12 min/week). This finding supports our hypothesis 1, that the web-based programs offered for the intervention and control group will enhance the PA in long term. However, this increase was not significantly different between the two tested groups. Furthermore, the effect size was marginal for the secondary outcomes (weight, eating behavior and quality of life) and not significant across groups. Thus, we cannot assume superiority of the IG, which can be interpreted that the non-interactive program containing evidence-based information (CG) was as effective as the interactive and tailored TK-FitnessCoach (IG) to increase PA. This result does not

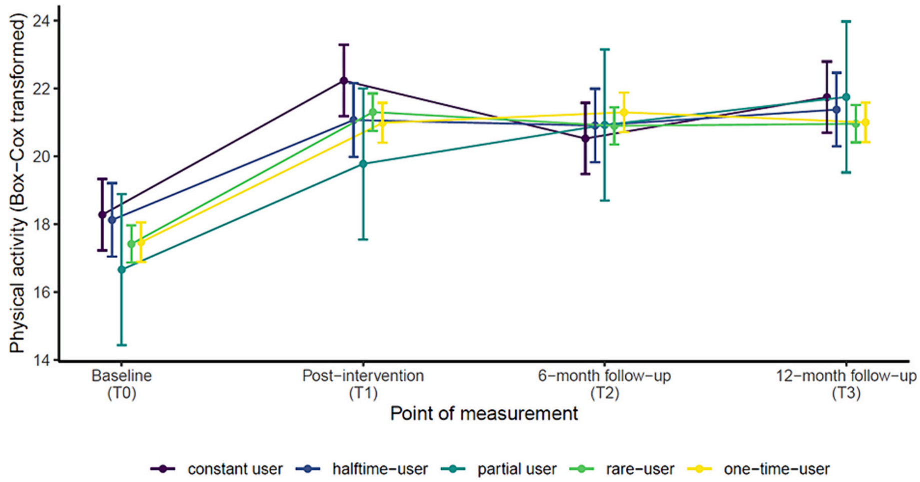

Fig. 3. Change in PA for different user types. User types based on weekly login activity. Error bars represents 95 % confidence intervals. Imputed ITT data set, only intervention group.
Sample distribution: constant user (n = 124); halftime-user (n = 116); partial user (n = 27); rare-user (n = 474); one-time-user (n = 412).

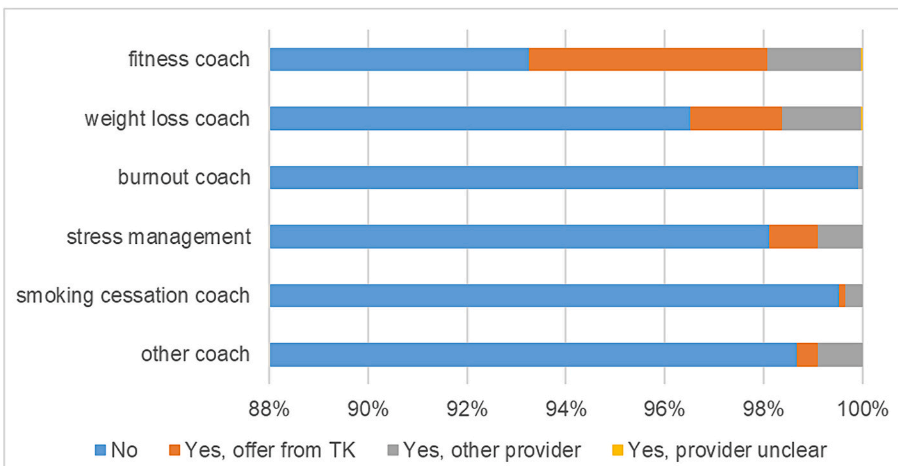

Fig. 4. Usage of other online health programs at T0 (relative shares). ITT data set.

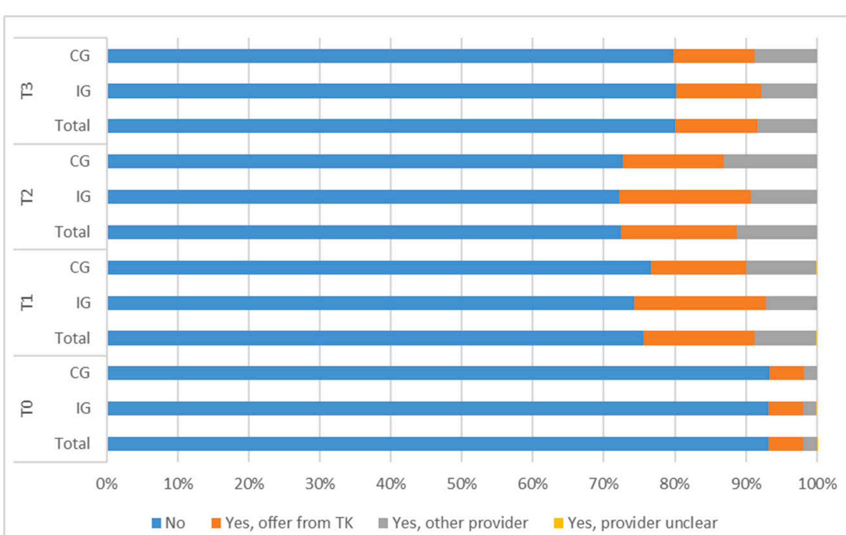

Fig. 5. Usage of other online fitness programs over time (relative shares) differentiated by CG and IG. ITT data set.

support hypothesis 2. Lastly, usage behavior was not associated with higher PA, which suggests rejecting hypothesis 3 as well.

4.2. Limitations

This study shows some limitations that need to be discussed. In our study, we implemented self-reported PA, which is discussed as an imprecise measurement (Martins et al., 2017). It is unclear whether our results might have changed if we would have recorded PA by fitness monitoring tools, taking into account, that the implementation of fitness trackers or accelerometers (integrated in common smartphones) might enhance PA (Kwan et al., 2020; Oosterveen et al., 2017; Foster et al., 2013; Stamatakis et al., 2022). Nevertheless, fitness trackers were used in the medical sub study of the TK FitnessCoach. There, comparable time effects on PA were found, but differences between IG and CG were also not detectable (Kohl et al., 2022). As consequence, we assume that a potential increase in measurement precision by implementing fitness trackers or accelerometers would not have changed our results. Furthermore, the application of fitness trackers or accelerometers would have biased our results in another way since we cannot imply that the common basic population uses those technical devices in a real world setting, although their use is increasing within the last years. Accelerometers have been found to be suitable instruments to reliably measure objective PA. However, a systematic review concluded that especially in the case of light physical activity, accelerometers show less accuracy (Lynch et al., 2019).

A general difficulty of self-reported outcomes like PA measurement might be that respondents needed to report their current PA quite spontaneously. Thus, the baseline measurement might be less valid than the subsequent measurements, since respondents are aware of stronger self-monitoring during the course of the intervention.

Another limiting factor is related to the measurement time points. Though we consequently applied reminders and excluded participants if they did not answer to a questionnaire within 8 weeks, this study design enables a possible variance in measurement time. Technically, it was possible that responses differ up to 8 weeks between individuals within one time point. Thus, some late responding individuals would have 2 months longer to spend time on PA, which might affect the outcome. However, we do not see any patterns that might systematically bias the response behavior and therefore accept this as lack of precision in measurement. Furthermore, the primary outcome uses a reference period of 4 weeks, ensuring that the measurement is consistent across individuals regardless of a delayed response.

A further point that needs to be discussed is the coverage of our sample. As we described, the ITT sample consists of 76 % female participants, 72 % are living in a stable partnership (i.a. married), and about 76 % are employed. It stands to question, whether specific vulnerable groups, e.g. unemployed persons had higher barriers to access the RCT and thus are undercovered in our sample. From statistical perspective, this pattern would has as consequence a limited validity of our study in terms of coverage of the sample. On the other hand, this approximately reflects the use of prevention offers (Lahtio et al., 2022). Our data represent a realistic and thus representative setting of the potential users of the TK-FitnessCoach. From the perspective of the coach provider, it would be more difficult to reach men, singles or people who are not employed to convince them to use the health coach.

A closer look to the data showed that those participants with higher PA at T0 decreased their mean time spent on PA until T1 and remained constant afterwards. For those, who showed low PA at T0, a stronger increase until T1 was found and their PA remained relatively constant afterwards, too. This pattern indicates a regression to the mean. Regardless of IG or CG, for users, starting on a high level of time spent on PA, it might be more difficult to maintain this high level in the long term, whereas, for users starting from a low level of time spent on PA, it might be rather easy to increase their PA. Further, it is assumable that the majority of users in both groups stabilize their time spent for PA, whether at a higher or a lower level.

In our ITT sample, we had to impute a huge amount of cases, which might limit the informative value of our results. Since we had many dropouts over time, the MAR assumption might be violated. For sensitivity analyses, we based the models on the complete cases. This approach, however, might overestimate the true effect, since participants might systematically drop out of the study due to study characteristics. Therefore, we assume that the true effect of the TK-FitnessCoach lies between the complete cases estimator and the ITT estimator, which were found to be quite similar in our analyses.

The dose-response-relationship was modelled in our study by including different usage behaviors and classifying them into user types. This transformation comes along with a certain information loss. However, the advantage of this approach is that is does not only consider the login frequency but also the period of using the coach.

4.3. Strengths

Our study exhibits several strengths. Firstly, it encompasses a large sample size of the German population, whereas other methodologically similar studies utilized considerably smaller sample sizes (Grey et al., 2019; Wijsman et al., 2013; Vroege et al., 2014; Cavallo et al., 2012; Peels et al., 2013). Moreover, employing the randomized controlled trial (RCT) design, regarded as the highest level of evidence (Burns et al., 2011), allowed for the integration of modern analytical methods like multilevel modeling and multiple imputation. Additionally, our research extends the existing literature by examining long-term effectiveness over a broader timeframe compared to studies with observation periods ranging from 12 to 24 weeks (Wijsman et al., 2013; Vroege et al., 2014; Cavallo et al., 2012; Peels et al., 2013).

Furthermore, our RCT study implemented a quasi-intervention as control group. A typology study concluded that the choice of control group has a high impact on effect sizes observed in RCTs that are focusing on the evaluation of mHealth interventions (Goldberg et al., 2023). The authors suggest a five class scheme with comparative strength ranging from low (class 5) to high (class 1). According to this scheme, our RCT can be categorized as additive design, testing an intervention effect (control group) with an active component added in the intervention group. By this, it represents a study of the category I with high comparative power.

4.4. Comparison with prior work

Due to ethical reasons and the impossibility to control human behavior in a real world scenario, all participants might have used any information available to increase their fitness. This is partly represented by our exploratory analyses. 20–28 % of the respondents of both groups reported using other online fitness programs. Whether participation in this study generated an increasing interest of participants in other online health or fitness programs to increase their fitness or if the enhanced use of health programs reflected the general use of digital technology, especially with the beginning of the COVID-19 pandemic remains unclear. In a population survey from 2020 with 1014 internet users in Germany, 22.7 % of persons reported using digital technology to support themselves in moderate PA (De Santis et al., 2021). Since our analyses indicated that there is no difference between IG and CG in this usage behavior, we can though assume that this has no effect on the investigated difference between the groups. However, this pattern might in general confound our results.

The COVID-19 pandemic could have influenced the outcomes of this research. The enforced restrictions on indoor physical activities could have altered the subjects' exercise habits within the bounds of legal regulations. Due to the ever-changing nature of the pandemic, we were unable to model potential pandemic effects into our models.

Apart from the pandemic, there could have been other seasonal influences. During the winter months, which coincided with the measurement time point T2 for many subjects, outdoor activities for physical exercise may have been less appealing (Wagner et al., 2019). However, as both the IG and the CG would have been affected by this, it would only account for a general decrease in PA for both groups, without causing a difference between the groups. Winter months might also be related to the higher usage of additional online fitness programs at T2 in both groups, as this might be true for the period between T0 and T1, the beginning of the COVID-19 pandemic.

A further point that needs to be discussed is whether the found increase in PA is of clinical relevance. The WHO and other medical associations recommend moderate-intensity aerobic PA for at least 150 min/week for health prevention (World Health Organization, 2020). The participants of this study started with a high level of total mean PA of >270 min/week, which increased to 356 min/week at T1. Even at T3, the level of PA remained high for both groups (IG: 351 min/week; CG 347 min/week). However, the high and increasing values of the standard deviation of general PA (IG: ± 248.76) and especially sports activity (IG: ±124.86), latter exceeding the mean value at T0, reflect extreme differences in the PA of participants from the beginning of the study until its end. This shows the heterogeneity of the participants in our study. Although we cannot differentiate clearly the PA according to its intensity (low, moderate, intensive, aerobic, anaerobic), the general increase in PA can be considered to be a positive signal. Various studies found that even a moderate increase in PA, starting from a far lower level than 150 min/week has positive effects on health outcomes, e.g., on reducing hypertension, diabetes, arteriosclerosis, cardiovascular diseases, and all-cause mortality (Ahmadi et al., 2022; Hansen et al., 2020; Paluch et al., 2023). According to a U.S. population study published in 2020, the highest reduction of all-cause mortality was detected in groups that doubled their PA from 4000 to 8000 steps per day (adjusted hazard ratio: 0.49; CI: [0.44–0.55]) (Saint-Maurice et al., 2020). Higher step increases were beneficial until 12,000 steps. Even lower increases, starting at a level of 2000 steps, were considered advantageous. This finding would be encouraging for persons with low general PA.

A study investigating behavior changes during the COVID-19 pandemic in Spring 2020 in Germany found that 18.5 % of the population increased their PA, while 29.4 % decreased it. People who increased PA were more often female and had higher education (Klosterhalfen et al., 2022). In our study, 75 % of participants was female, 68 % had a university entrance qualification but we did not find associations between PA and gender or PA or educational level. We assume that only a low proportion of persons with a higher need for PA increase, due to health-related risk behavior, were reached by the TK-FitnessCoach study. Those from this target group who nonetheless participated in this study were rarely successful in increasing their PA.

Our finding, that both the CG and the IG showed an increase in PA related to the intervention is not uncommon. A systematic review states that 28 % of 29 studies reviewed report significant improvement in PA for their control groups (Waters et al., 2012). Thus, a possible explanation for the found increase in PA might be related to the general participation in the study as a motivator that might strengthen behavior change.

In line with an increase in PA for both groups, we observed a significant but marginal increase in healthy eating behavior over time that, according to the validation study of this instrument might have correlated with BMI reduction (Wurst et al., 2022)] The effects of healthy eating behavior were not significantly different between the groups. We did not examine if participants in the IG used the offered additional interactive modules targeting healthy diet. However, these findings suggest that the participation in the study aiming to increase fitness might have a small beneficial effect on weight- and health-conscious diet in both groups. Interestingly, the concurrent RCT focusing on the health goal “weight reduction and maintaining weight” did not find any effects neither on weight nor on healthy eating (Fichtner et al., 2022). This is supposed to be another hint that at least a part of the participants in this study started with high motivation to increase health behavior change and was able to maintain it.

Sedentary behavior has a negative impact on physical and mental health, but can be at least partially compensated by physical activity (Nyström et al., 2019) (De Mello et al., 2013) (Alfaddagh et al., 2020). In our study we found no positive effect of physical activity on mental or physical health. However, as health impairments were an exclusion criterion in our study, the symptom burden in our RCT was lower compared to studies aiming to improve physical or mental outcomes of people with diseases. Nevertheless, it has to be considered that in our study, we did not measure sedentary behavior since the focus was on increasing physical activity measured by active mins/week. However, the integration of a sedentary behavior measurement might have contributed to a better understanding of the effectiveness of the TK-HealthCoach.

5. Conclusions

Based on our study, we could draw some conclusions. First, web-based health programs have the potential to significantly increase PA among adults. This is consistent across various self-reported measurements of PA. Second, no difference was found between the interactive and the non-interactive version of the TK-FitnessCoach, suggesting that the mere provision of PA-related information, whether through interactive or non-interactive means, can be sufficient to motivate adults to increase their PA. However, the results must be interpreted with caution, taking into account possible limited validity of self-reported measurement of PA, a general study participation effect, and limited accessibility of low-motivated persons. Third, the steep increase in PA between T0 and T1 and the maintenance of a high level until T3 is positive in terms of the sustainability of the program. However, this positive outcome only applies to a part of the study population, as shown by the extreme difference in PA, represented by a high standard deviation between the individuals at each measurement point. Fourth, since no dose-response-relationship was found, we conclude that the amount of engagement with the interactive program does not directly correlate with the magnitude of increase in PA. As the PA increased the same in the CG, it is assumable that high motivation at baseline, maybe triggered by the COVID-19 pandemic, indicates that even minimal use of the TK-FitnessCoach or well-presented information on a static website could be beneficial. Fifth, we found also small positive effects on eating behavior in both groups that is associated with a healthy lifestyle. Sixth, there was no evidence of potentially undetected differences between IG and CG due to the additional use of other fitness programs. Finally, since our sample showed already high PA at T0, we conclude that persons with higher PA are more likely to benefit from participation in any fitness program offered than persons with a low baseline level of PA.

For the improvement of health prevention in a broad population, the benefits of even small increases in exercise should be consistently communicated and integrated into online fitness programs. Health campaigns should be launched and different ways to access affected persons should be improved, e.g., via general practitioners, occupational health care, health insurance companies, and the involvement of local authorities, to support people with low levels of PA in increasing their motivation to change their health behavior. Online programs of different types, in combination with selectable human contact or avatars to achieve more tailored counseling, could be a promising approach that should be investigated in future studies.

Both groups received a comprehensive collection of informative materials and it appears that in this study, interactive elements do not have advantages regarding PA or other health outcomes in comparison to well-presented written information. It remains unclear which components in digital interventions contribute significantly to an increase in PA.

Abbreviations

BSA-F	Bewegungs- und Sportaktivität Fragebogen
CC	complete cases
CG	Control Group
CI	Confidence Interval
DALYs	disability-adjusted life years
GLMM	generalized linear mixed models
hc-EB	health-conscious eating behavior
IG	Intervention Group
ITT	Intention to Treat
MAR	Missing at Random
min/week	Minutes per Week
MTU	Multi-time-user
PA	Physical Activity
QoL	Quality of Life
RCT	Randomized Controlled Trial
RIV	Relative Increase in Variance
TK	Techniker Krankenkasse
wc-EB	weight-controlling eating behavior
WHO	World Health Organization

Declaration of competing interest

The two involved departments Section for Health Services Research and Rehabilitation Research, (SEVERA) at the Institute of Medical Biometry and Statistics and the Department of Sport and Sport were commissioned by the Techniker Krankenkasse (German Health Insurance Company) for the scientific evaluation of the web-based health program. All authors report funding from the Techniker Krankenkasse for clinical trial design, implementation and scientific evaluation within this study. Apart from funding, the Techniker Krankenkasse had no influence on the evaluation of this study. All authors declare that they have no conflicts of interest.

Acknowledgements

This project was funded by the Techniker Krankenkasse (German Health Insurance Company). The project funder had no influence on the planning and implementation of the study, the analysis and interpretation of the data, or the publication of the results. We thank all study participants, who supported our study. Furthermore, we would like to thank Techniker Krankenkasse, represented by Dr. Nicole Knaack, Dr. Kerstin Hofreuter-Gätgens, and Dagmar Köppel, for funding this study. Additionally, our gratitude extends to Irina Kopman and her team at Vilua Healthcare GmbH for IT assistance. We also acknowledge support by the Open Access Publication Fund of the University of Freiburg.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.invent.2024.100762.

References

Ahmadi, M.N., Gill, J.M.R., Stamatakis, E., 2022. Association of changes in physical activity and adiposity with mortality and incidence of cardiovascular disease: longitudinal findings from the UK Biobank. Mayo Clin. Proc. 97 (5), 847–861 (PMID: 35410749).

Alfaddagh, A., Martin, S.S., Leucker, T.M., Michos, E.D., Blaha, M.J., Lowenstein, C.J., Jones, S.R., Toth, P.P., 2020. Inflammation and cardiovascular disease: from mechanisms to therapeutics. Am. J. Prev. Cardiol. 4, 100130 https://doi.org/10.1016/j.ajpc.2020.100130.

Audigier, V., Resche-Rigon, M., 2021. micemd: Multiple Imputation by Chained Equations With Multilevel Data (Version 1.8.0). Available from: https://CRAN.R-project.org/package=micemd.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67 (1) https://doi.org/10.18637/jss.v067.i01.

Ben-Shachar, M., Lüdecke, D., Makowski, D., 2020. Effectsize: estimation of effect size indices and standardized parameters. JOSS 5 (56), 2815. https://doi.org/10.21105/joss.02815.

Blodgett, J.M., Mitchell, J.J., Stamatakis, E., Chastin, S., Hamer, M., 2023. Associations between the composition of daily time spent in physical activity, sedentary behaviour and sleep and risk of depression: compositional data analyses of the 1970 British cohort Study. J. Affect. Disord. 320, 616–620 (PMID:36183826).

Brame, J., Kohl, J., Wurst, R., Fuchs, R., Tinsel, I., Maiwald, P., Fichtner, U., Armbruster, C., Bischoff, M., Farin-Glattacker, E., Lindinger, P., Bredenkamp, R., Gollhofer, A., König, D., 2022. Health effects of a 12-week web-based lifestyle intervention for physically inactive and overweight or obese adults: study protocol of two randomized controlled clinical trials. IJERPH 19 (3), 1393. https://doi.org/10.3390/ijerph19031393.

Brame, J., Kohl, J., Centner, C., Wurst, R., Fuchs, R., Tinsel, I., Maiwald, P., Fichtner, U. A., Sehlbrede, M., Farin-Glattacker, E., Gollhofer, A., König, D., 2023. Effects of a web-based lifestyle intervention on physical fitness and health in physically inactive adults: a randomized controlled trial. Healthcare 11 (21), 2847. https://doi.org/10.3390/healthcare11212847.

Brooks, M.E., Kristensen, K., Benthem, K.J., van Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Mächler, M., Bolker, B.M., 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R J. 9 (2), 378. https://doi.org/10.32614/RJ-2017-066.

Burns, P.B., Rohrich, R.J., Chung, K.C., 2011. The levels of evidence and their role in evidence-based medicine. Plast. Reconstr. Surg. 128 (1), 305–310. https://doi.org/10.1097/PRS.0b013e318219c171.

van Buuren, S., 2018. Flexible Imputation of Missing Data, Second Edition. 2nd ed. Second edition. | Boca Raton, Florida : CRC Press, [2019] |: Chapman and Hall/CRC. https://doi.org/10.1201/9780429492259ISBN:978-0-429-49225-9.

Cavallo, D.N., Tate, D.F., Ries, A.V., Brown, J.D., DeVellis, R.F., Ammerman, A.S., 2012. A social media–based physical activity intervention. Am. J. Prev. Med. 43 (5), 527–532. https://doi.org/10.1016/j.amepre.2012.07.019.

Chen, L.-J., Hamer, M., Lai, Y.-J., Huang, B.-H., Ku, P.-W., Stamatakis, E., 2022. Can physical activity eliminate the mortality risk associated with poor sleep? A 15-year follow-up of 341,248 MJ Cohort participants. J. Sport Health Sci. 11 (5), 596–604 (PMID:33713846).

Davies, C., Corry, K., Van Itallie, A., Vandelanotte, C., Caperchione, C., Mummery, W.K., 2012. Prospective associations between intervention components and website engagement in a publicly available physical activity website: the case of 10,000 steps Australia. J. Med. Internet Res. 14 (1), e4 (PMID:22260810).

De Mello, M.T., Lemos, V. de A., HKM, Antunes, Bittencourt, L., Santos-Silva, R., Tufik, S., 2013. Relationship between physical activity and depression and anxiety symptoms: a population study. J. Affect. Disord. 149 (1–3), 241–246. https://doi.org/10.1016/j.jad.2013.01.035.

De Santis, K.K., Jahnel, T., Sina, E., Wienert, J., Zeeb, H., 2021. Digitization and health in Germany: cross-sectional nationwide survey. JMIR Public Health Surveill. 7 (11), e32951 (PMID:34813493).

Enders, C.K., 2010. Applied Missing Data Analysis. Guilford Press, New York (ISBN:978-1-60623-639-0).

Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 39 (2), 175–191. https://doi.org/10.3758/BF03193146.

Fichtner, U.A., Armbruster, C., Bischoff, M., Maiwald, P., Sehlbrede, M., Tinsel, I., Brame, J., Kohl, J., König, D., Fuchs, R., Wurst, R., Farin-Glattacker, E., 2022. Evaluation of an interactive web-based health program for weight loss—a randomized controlled trial. IJERPH 19 (22), 15157. https://doi.org/10.3390/ijerph192215157.

Finger, J.D., Mensink, G.B.M., Lange, C., Manz, K., 2017. Gesundheitsfördernde körperliche Aktivität in der Freizeit bei Erwachsenen in Deutschland. J. Health Monit. 2 (2).

Foster, C., Richards, J., Thorogood, M., Hillsdon, M., 2013. Remote and web 2.0 interventions for promoting physical activity. Cochrane Database Syst. Rev. 9 (9), CD010395 (PMID:24085594).

Fuchs, R., 1997. Psychologie und körperliche Bewegung. Hogrefe, Grundlagen für theoriegeleitete Interventionen.

Fuchs, R., Göhner, W., Mahler, C., Fleitz, A., Seelig, H., 2009. Theoriegeleitete Lebensstiländerung. Centaurus-Verl.: Freiburg im Breisgau. https://doi.org/10.25656/01:2249.

Fuchs, R., Klaperski, S., Gerber, M., Seelig, H., 2015. Messung der Bewegungs- und Sportaktivität mit dem BSA-Fragebogen: Eine methodische Zwischenbilanz. Zeitschrift für Gesundheitspsychologie 23 (2), 60–76. https://doi.org/10.1026/0943-8149/a000137.

GBD 2019 Risk Factors Collaborators, 2020. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 396 (10258), 1223–1249 (PMID:33069327).

Glattacker, M., Meixner, K., Farin, E., Jäckel, W.H., 2007. Entwicklung eines rehabilitationsspezifischen Komorbiditätsscores und erste Prüfung methodischer Gütekriterien. Phys. Rehab. Kur. Med. 17 (5), 260–270. https://doi.org/10.1055/s-2007-970405.

Goldberg, S.B., Sun, S., Carlbring, P., Torous, J., 2023. Selecting and describing control conditions in mobile health randomized controlled trials: a proposed typology. npj Digit. Med. 6 (1), 181. https://doi.org/10.1038/s41746-023-00923-7.

Grey, E.B., Thompson, D., Gillison, F.B., 2019. Effects of a web-based, evolutionary mismatch-framed intervention targeting physical activity and diet: a randomised controlled trial. Int. J. Behav. Med. 26 (6), 645–657. https://doi.org/10.1007/s12529-019-09821-3.

Grund, S., Robitzsch, A., Luedtke, O., 2019. mitml: Tools for Multiple Imputation in Multilevel Modeling (Version 0.3-7). Available from: https://cran.r-project.org/web/packages/mitml/mitml.pdf.

Hansen, B.H., Dalene, K.E., Ekelund, U., Wang Fagerland, M., Kolle, E., Steene-Johannessen, J., Tarp, J., Alfred, Anderssen S., 2020. Step by step: association of device-measured daily steps with all-cause mortality-A prospective cohort study. Scand. J. Med. Sci. Sports 30 (9), 1705–1711 (PMID:32427398).

Hartig, F., 2022. DHARMa: Residual Diagnostic for Hierarchical (Multi-Level / Mixed) Regression Models. Available from: https://cran.r-project.org/web/packages/DHARMa/vignettes/DHARMa.html.

Jahangiry, L., Farhangi, M.A., Shab-Bidar, S., Rezaei, F., Pashaei, T., 2017. Web-based physical activity interventions: a systematic review and meta-analysis of randomized controlled trials. Public Health 152, 36–46 (PMID:28734170).

Jurik, R., Stastny, P., 2019. Role of nutrition and exercise programs in reducing blood pressure: a systematic review. J. Clin. Med. 8 (9), 1393 (PMID:31492032).

Klosterhalfen, S., Kotz, D., Kastaun, S., 2022. Did smoking, alcohol consumption, and physical activity change during the COVID-19 restrictions in Germany in spring 2020?: findings from a cross-sectional population survey (the DEBRA study). SUCHT 68 (3), 129–138. https://doi.org/10.1024/0939-5911/a000765.

Kohl, J., Brame, J., Hauff, P., Wurst, R., Sehlbrede, M., Fichtner, U.A., Armbruster, C., Tinsel, I., Maiwald, P., Farin-Glattacker, E., Fuchs, R., Gollhofer, A., König, D., 2022. Effects of a web-based weight loss program on the healthy eating index-NVS in adults with overweight or obesity and the association with dietary, anthropometric and cardiometabolic variables: a randomized controlled clinical trial. Nutrients 15 (1), 7. https://doi.org/10.3390/nu15010007.

Kohl, J., Brame, J., Centner, C., Wurst, R., Fuchs, R., Sehlbrede, M., Tinsel, I., Maiwald, P., Fichtner, U.A., Armbruster, C., Farin-Glattacker, E., Gollhofer, A., König, D., 2023. Effects of a web-based lifestyle intervention on weight loss and Cardiometabolic risk factors in adults with overweight and obesity: randomized controlled clinical trial. J. Med. Internet Res. 25, e43426 (PMID:37368484).

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017. lmerTest package: tests in linear mixed effects models. J. Stat. Softw. 82 (13) https://doi.org/10.18637/jss.v082.i13.

Kwan, R.Y.C., Salihu, D., Lee, P.H., Tse, M., Cheung, D.S.K., Roopsawang, I., Choi, K.S., 2020. The effect of e-health interventions promoting physical activity in older people: a systematic review and meta-analysis. Eur. Rev. Aging Phys. Act. 17, 7 (PMID:32336996).

Lahtio, H., Rintala, A., Immonen, J., Sjögren, T., 2022. The effectiveness of physical activity-promoting web- and mobile-based distance weight loss interventions on body composition in rehabilitation settings: systematic review, meta-analysis, and meta-regression analysis. J. Med. Internet Res. 24 (3), e25906 https://doi.org/10.2196/25906.

Lenth, R., 2021. emmeans: Estimated Marginals Means, aka Least-Squares Means (Version 1.7.1-1) [Software]. Available from: https://CRAN.R-project.org/package=emmeans.

Löwe, B., Spitzer, R.L., Zipfel, S., Herzog, W.. Autorisierte deutsche Version des "Prime MD Patient Health Questionnaire (PHQ)". Available from: https://www.klinikum.uni-heidelberg.de/fileadmin/Psychosomatische_Klinik/download/PHQ_Manual1.pdf.

Lynch, B.A., Kaufman, T.K., Rajjo, T.I., Mohammed, K., Kumar, S., Murad, M.H., Gentile, N.E., Koepp, G.A., McCrady-Spitzer, S.K., Levine, J.A., 2019. Accuracy of accelerometers for measuring physical activity and levels of sedentary behavior in children: a systematic review. J. Prim. Care Community Health 10, 215013271987425. https://doi.org/10.1177/2150132719874252.

Maiwald, P., Bischoff, M., Lindinger, P., Tinsel, I., Sehlbrede, M., Fichtner, U.A., Metzner, G., Schlett, C., Farin-Glattacker, E., 2023. The effect of interactivity, tailoring, and use intensity on the effectiveness of an internet-based smoking cessation intervention over a 12-month period: randomized controlled trial. J. Med. Internet Res. 25, e47463 https://doi.org/10.2196/47463.

Martins, J.C., Aguiar, L.T., Nadeau, S., Scianni, A.A., Teixeira-Salmela, L.F., Faria, C.D.C. de M., 2017. Measurement properties of self-report physical activity assessment tools in stroke: a protocol for a systematic review. BMJ Open 7 (2), e012655. https://doi.org/10.1136/bmjopen-2016-012655.

McKnight, P.E. (Ed.), 2007. Missing Data: A Gentle Introduction. Guilford Press, New York (ISBN:978-1-59385-393-8).

Mclaughlin, M., Delaney, T., Hall, A., Byaruhanga, J., Mackie, P., Grady, A., Reilly, K., Campbell, E., Sutherland, R., Wiggers, J., Wolfenden, L., 2021. Associations between digital health intervention engagement, physical activity, and sedentary behavior: systematic review and meta-analysis. J. Med. Internet Res. 23 (2), e23180 (PMID:33605897).

Nakagawa, S., Johnson, P.C.D., Schielzeth, H., 2017. The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. J. R. Soc. Interface 14 (134), 20170213. https://doi.org/10.1098/rsif.2017.0213.

Nyström, M.B.T., Hassmén, P., Sörman, D.E., Wigforss, T., Andersson, G., Carlbring, P., 2019. Are physical activity and sedentary behavior related to depression? Martino G, editor. Cogent Psychol. 6 (1), 1633810 https://doi.org/10.1080/23311908.2019.1633810.

Oosterveen, E., Tzelepis, F., Ashton, L., Hutchesson, M.J., 2017. A systematic review of eHealth behavioral interventions targeting smoking, nutrition, alcohol, physical activity and/or obesity for young adults. Prev. Med. 99, 197–206 (PMID:28130046).

Pahlke, F., König, I.R., Ziegler, A., 2004. Randomization In Treatment Arms (RITA): Ein Randomisierungs-Programm für klinische Studien. Informatik Biometrie und Epidemiologie in Medizin und Biologie 35 (1), 1–22.

Paluch, A.E., Bajpai, S., Ballin, M., Bassett, D.R., Buford, T.W., Carnethon, M.R., Chernofsky, A., Dooley, E.E., Ekelund, U., Evenson, K.R., Galuska, D.A., Jefferis, B.J., Kong, L., Kraus, W.E., Larson, M.G., Lee, I.-M., Matthews, C.E., Newton, R.L., Nordström, A., Nordström, P., Palta, P., Patel, A.V., Pettee Gabriel, K., Pieper, C.F., Pompeii, L., Rees-Punia, E., Spartano, N.L., Vasan, R.S., Whincup, P.H., Yang, S., Fulton, J.E., 2023. Steps for health collaborative. Prospective association of daily steps with cardiovascular disease: a harmonized meta-analysis. Circulation 147 (2), 122–131 (PMID:36537288).

Peels, D., Bolman, C., Golsteijn, R.H.J., de Vries, H., Mudde, A., van Stralen, M., Lechner, L., 2013. Long-term efficacy of a printed or a web-based tailored physical activity intervention among older adults. Int. J. Behav. Nutr. Phys. Act. 10 (1), 104. https://doi.org/10.1186/1479-5868-10-104.

Pelliccia, A., Sharma, S., Gati, S., Bäck, M., Börjesson, M., Caselli, S., Collet, J.-P., Corrado, D., Drezner, J.A., Halle, M., Hansen, D., Heidbuchel, H., Myers, J., Niebauer, J., Papadakis, M., Piepoli, M.F., Prescott, E., Roos-Hesselink, J.W., Graham Stuart, A., Taylor, R.S., Thompson, P.D., Tiberi, M., Vanhees, L., Wilhelm, M., ESC Scientific Document Group, 2021. 2020 ESC Guidelines on sports cardiology and exercise in patients with cardiovascular disease. Eur. Heart J. 42 (1), 17–96. 32860412.

Posadzki, P., Pieper, D., Bajpai, R., Makaruk, H., Könsgen, N., Neuhaus, A.L., Semwal, M., 2020. Exercise/physical activity and health outcomes: an overview of Cochrane systematic reviews. BMC Public Health 20 (1), 1724 (PMID:33198717).

Renner, B., Hahn, A., Schwarzer, R., 1996. Freie Universität Berlin, editors. Risiko und Gesundheitsverhalten: Dokumentation der Meßinstrumente des Forschungsprojekts "Berlin Risk Appraisal and Health Motivation Study" (BRAHMS). Freie Univ. Berlin, Berlin (ISBN:978-3-00-000631-9).

Resche-Rigon, M., White, I.R., 2018. Multiple imputation by chained equations for systematically and sporadically missing multilevel data. Stat. Methods Med. Res. 27 (6), 1634–1649. https://doi.org/10.1177/0962280216666564.

Saint-Maurice, P.F., Troiano, R.P., Bassett, D.R., Graubard, B.I., Carlson, S.A., Shiroma, E.J., Fulton, J.E., Matthews, C.E., 2020. Association of daily step count and step intensity with mortality among US adults. JAMA 323 (12), 1151–1160 (PMID:32207799).

Schuch, F., Vancampfort, D., Firth, J., Rosenbaum, S., Ward, P., Reichert, T., Bagatini, N.C., Bgeginski, R., Stubbs, B., 2017. Physical activity and sedentary behavior in people with major depressive disorder: A systematic review and meta-analysis. J. Affect. Disord. 210, 139–150 (PMID:28033521).

Sehlbrede, M., Wurst, R., 2022. fitevalapp. Available from: https://github.com/Methodicum/fitevalapp.

Stamatakis, E., Ahmadi, M.N., Gill, J.M.R., Thøgersen-Ntoumani, C., Gibala, M.J., Doherty, A., Hamer, M., 2022. Association of wearable device-measured vigorous intermittent lifestyle physical activity with mortality. Nat. Med. 28 (12), 2521–2529 (PMID:36482104).

Tinsel, I., Metzner, G., Schlett, C., Sehlbrede, M., Bischoff, M., Anger, R., Brame, J., König, D., Wurst, R., Fuchs, R., Lindinger, P., Bredenkamp, R., Farin-Glattacker, E., 2021. Effectiveness of an interactive web-based health program for adults: a study protocol for three concurrent controlled-randomized trials (EVA-TK-Coach). Trials 22 (1), 526. https://doi.org/10.1186/s13063-021-05470-8.

Vroege, D.P., Wijsman, C.A., Broekhuizen, K., de Craen, A.J.M., van Heemst, D., van der Ouderaa, F.J.G., van Mechelen, W., Slagboom, P.E., Catt, M., Westendorp, R.G.J., Verhagen, E.A.L.M., Mooijaart, S.P., 2014. Dose-response effects of a web-based physical activity program on body composition and metabolic health in inactive older adults: additional analyses of a randomized controlled trial. J. Med. Internet Res. 16 (12), e265 (PMID:25486673).

Wagner, A.L., Keusch, F., Yan, T., Clarke, P.J., 2019. The impact of weather on summer and winter exercise behaviors. J. Sport Health Sci. 8 (1), 39–45. https://doi.org/10.1016/j.jshs.2016.07.007.

Waters, L., Reeves, M., Fjeldsoe, B., Eakin, E., 2012. Control group improvements in physical activity intervention trials and possible explanatory factors: a systematic review. J. Phys. Act. Health 9 (6), 884–895. https://doi.org/10.1123/jpah.9.6.884.

Wen, H., Wang, L., 2017. Reducing effect of aerobic exercise on blood pressure of essential hypertensive patients: a meta-analysis. Medicine (Baltimore) 96 (11), e6150 (PMID:28296729).

Wickham, H., 2016. ggplot2. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-24277-4ISBN:978-3-319-24275-0.

Wijsman, C.A., Westendorp, R.G., Verhagen, E.A., Catt, M., Slagboom, P.E., de Craen, A.J., Broekhuizen, K., van Mechelen, W., van Heemst, D., van der Ouderaa, F., Mooijaart, S.P., 2013. Effects of a web-based intervention on physical activity and metabolism in older adults: randomized controlled trial. J. Med. Internet Res. 15 (11), e233 https://doi.org/10.2196/jmir.2843.

Wirtz, M.A., Morfeld, M., Glaesmer, H., Brähler, E., 2018 Apr. Konfirmatorische Prüfung der Skalenstruktur des SF-12 Version 2.0 in einer deutschen bevölkerungs-repräsentativen Stichprobe. Diagnostica 64 (2), 84–96. https://doi.org/10.1026/0012-1924/a000194.

World Health Organization, 2020. WHO guidelines on physical activity and sedentary behaviour: at a glance. Available from: https://iris.who.int/bitstream/handle/10665/337001/9789240014886-eng.pdf?sequence=1.

Wurst, R., Brame, J., Ramsenthaler, C., König, D., Fuchs, R., 2022. A questionnaire to assess eating behavior: structure, validity and responsiveness of a new German eating behavior scale (SEV). Appetite 168, 105668. https://doi.org/10.1016/j.appet.2021.105668.