— **Post-print version** —

Article accepted in
Journal of Medical Internet Research

Please cite as:

Mazeas A, Duclos M, Pereira B, Chalabaev A. (2022). Evaluating the Effectiveness of Gamification on Physical Activity: Systematic Review and Meta-analysis of Randomized Controlled Trials. J Med Internet Res 2022;24(1):e26779. doi: [10.2196/26779](https://doi.org/10.2196/26779). PMID: 34982715

Does gamification improve physical activity? A systematic review and meta-analysis

Alexandre Mazeas, Martine Duclos, Bruno Pereira, Aïna Chalabaev

Submitted to: Journal of Medical Internet Research
on: December 25, 2020

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5
Supplementary Files..... 40
 Multimedia Appendixes 41
 Multimedia Appendix 1..... 41

Does gamification improve physical activity? A systematic review and meta-analysis

Alexandre Mazeas^{1, 2, 3}; Martine Duclos^{2, 4}; Bruno Pereira⁵; Aïna Chalabaev¹

¹SENS Univ. Grenoble Alpes Saint Martin d'Hères FR

²INRAE, UNH, CRNH Auvergne Clermont Auvergne University Clermont-Ferrand FR

³Kiplin Nantes FR

⁴Department of Sport Medicine and Functional Exploration University Hospital CHU G. Montpied Clermont-Ferrand FR

⁵Department of Biostatistics unit (DRCI) Clermont-Ferrand University Hospital Clermont-Ferrand FR

Corresponding Author:

Alexandre Mazeas

SENS

Univ. Grenoble Alpes

1741 rue de la Piscine

Saint Martin d'Hères

FR

Abstract

Background: The use of gamification to change behaviors and promote physical activity is a promising avenue to tackle the global physical inactivity pandemic and the current prevalence of chronic diseases. However, we do not have yet evidence of the effectiveness of gamified interventions with the existence of mixed results in the literature.

Objective: This systematic review and meta-analysis aimed to evaluate the effectiveness of gamified interventions and their healthcare potential by testing the scalability and sustainability of their influence on physical activity (PA) and sedentary behaviors (SB).

Methods: Five electronic databases (Pubmed, EMBASE, Scopus, Web of Science and CENTRAL) were searched for randomized controlled trials (RCTs) published since 2010. Eligibility criteria were based on components of PICO (participants, interventions, comparators, and outcomes). Studies were included when they used gamified interventions in daily life with an active or inactive control group, and when they assessed a PA or SB outcome. Studies were excluded when the intervention was based on an active video game using a gaming console. We conducted meta-analyses using a random effects model approach. Sensitivity analyses, influence analyses and publication bias analyses were performed to examine the robustness of our results.

Results: The main meta-analysis performed on 16 studies and 2407 participants, revealed a small to medium summary effect of gamified interventions on PA behavior ($g = 0.42$ (95%CI [0.14, 0.69])). No statistical difference was found between different subgroups (adults vs adolescents, healthy participants vs adults with chronic diseases) and no interaction effects with moderators like age, gender or BMI, suggesting good scalability of gamified interventions. Moreover, the effect appears better on the step count (MD = +1609.56 steps per day (95%CI [372.39, 2846.73])) than on moderate-to-vigorous PA ($g = 0.31$ (95%CI [-0.19, 0.80])). The long-term effect (measured with follow-up averaging 14 weeks after the end of the intervention) is more volatile with a very small to small effect ($g = 0.15$, [0.07, 0.23]).

Conclusions: This meta-analysis confirms that gamified interventions are promising to promote PA in various publics. However, if the effect persists after the end of the program, suggesting that it is not just a novelty effect due to the playful nature of gamification, it seems volatile with a small long-term effect. The integration of gamification in more global healthcare interventions could be a way to address this low sustainability. Future rigorous trials are needed to explore these perspectives.

(JMIR Preprints 25/12/2020:26779)

DOI: <https://doi.org/10.2196/preprints.26779>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ Please make my preprint PDF available to anyone at any time (recommended).

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org/preprint/26779>, my manuscript will be published in JMIR Publications.

Original Manuscript

Review

Does gamification improve physical activity? A systematic review and meta-analysis

Alexandre Mazéas Correspondent^{1,2,3}, Martine Duclos Co-Author^{2,4}, Bruno Pereira Co-Author⁵, Aïna Chalabaev Co-Author¹

¹*Univ. Grenoble Alpes, SENS, F-38000 Grenoble, France*

²*INRAE, UNH, CRNH Auvergne, Clermont Auvergne University, F-63000 Clermont-Ferrand, France*

³*Kiplin, F-44200 Nantes, France*

⁴*Department of Sport Medicine and Functional Exploration, University Hospital CHU G. Montpied, F-63000 Clermont-Ferrand, France*

⁵*Department of Biostatistics unit (DRCI), Clermont-Ferrand University Hospital, F-63000 Clermont-Ferrand, France*

ABSTRACT

Background: The use of gamification to change behaviors and promote physical activity is a promising avenue to tackle the global physical inactivity pandemic and the current prevalence of chronic diseases. However, we do not have yet evidence of the effectiveness of gamified interventions with the existence of mixed results in the literature.

Objective: This systematic review and meta-analysis aimed to evaluate the effectiveness of gamified interventions and their healthcare potential by testing the scalability and sustainability of their influence on physical activity (PA) and sedentary behaviors (SB).

Methods: Five electronic databases (Pubmed, EMBASE, Scopus, Web of Science and CENTRAL) were searched for randomized controlled trials (RCTs) published since 2010. Eligibility criteria were based on components of PICO (participants, interventions, comparators, and outcomes). Studies were included when they used gamified interventions in daily life with an active or inactive control group, and when they assessed a PA or SB outcome. Studies were excluded when the intervention was based on an active video game using a gaming console. We conducted meta-analyses using a random effects model approach. Sensitivity analyses, influence analyses and publication bias

analyses were performed to examine the robustness of our results.

Results: The main meta-analysis performed on 16 studies and 2407 participants, revealed a small to medium summary effect of gamified interventions on PA behavior ($g = 0.42$ (95%CI [0.14, 0.69])). No statistical difference was found between different subgroups (adults vs adolescents, healthy participants vs adults with chronic diseases) and no interaction effects with moderators like age, gender or BMI, suggesting good scalability of gamified interventions. Moreover, the effect appears better on the step count (MD = +1609.56 steps per day (95%CI [372.39, 2846.73])) than on moderate-to-vigorous PA ($g = 0.31$ (95%CI [-0.19, 0.80])). The long-term effect (measured with follow-up averaging 14 weeks after the end of the intervention) is more volatile with a very small to small effect ($g = 0.15$, [0.07, 0.23]).

Conclusion: This meta-analysis confirms that gamified interventions are promising to promote PA in various publics. However, if the effect persists after the end of the program, suggesting that it is not just a novelty effect due to the playful nature of gamification, it seems volatile with a small long-term effect. The integration of gamification in more global healthcare interventions could be a way to address this low sustainability. Future rigorous trials are needed to explore these perspectives.

PROSPERO registration number: CRD42020186882

Keywords: gamification; physical activity; behavior change; intervention; e-health; health behavior; systematic review; meta-analysis

INTRODUCTION

Background

Physical inactivity and sedentary behaviors (SB) are one of the leading risk factors for global mortality [1]. Each year, physical inactivity is responsible for more than five million deaths worldwide [2]. In contrast, regular physical activity (PA) prevents the risk of developing chronic diseases [3,4], limits their progression [5,6], and reduces early mortality [7]. In parallel, recent studies have highlighted the need to distinguish the effects of PA from SB, as their effects on health are independent. Indeed, there is a dose-response relationship between total sedentary time per day and overall mortality that is independent from PA participation [7]. Meta-analyses demonstrate that the risk of mortality in adults increases steadily with a sedentary lifestyle of more than 3 hours per day, and more significantly when this time exceeds 7 hours per day [8].

In this context, it is urgent to develop interventions that can effectively change both PA and SB. For this purpose, digital health interventions constitute a new opportunity to take care of patients, by involving them in their treatment in a dynamic and interactive way. A promising avenue to capitalize the efficacy of digital interventions is gamification. Gamification is defined as the use of game design elements in non-game contexts [9]. By integrating game mechanisms in interventions that are initially devoid of them, the purpose of gamification is to integrate into daily life the ingredients that make games enjoyable, in order to motivate participants to engage in PA [10]. The use of “motivational affordances” created by gamification can influence psychological (e.g. motivation, attitude, and enjoyment) and physical outcomes (e.g. physical capacities) [11] and appears therefore as a potentially powerful technique for behavior change.

By gamifying PA, participants are encouraged to move and walk to play, and this tends to make their activity more enjoyable and playful [12]. Unlike serious games and other traditional video games which require a dedicated time, location, and implementation [13], gamification techniques concern instead global PA in all aspects of daily life. This is made possible by mobile technologies and wearable devices that can track and collect daily activities in a continuous and online manner. This allows to intervene directly on the lifestyle of individuals without adding time and material constraints for the participants.

However, several literature reviews show inconsistent results concerning the use of gamification in behavioral interventions, with on the one hand some studies demonstrating positive effects, and on the other hand studies providing more mixed effects [12,14,15]. These reviews also emphasize the lack of high-quality studies and highlight the need for more rigorous trials that can isolate the impact of gamification (i.e., randomized controlled trials, RCTs). Importantly, Koivisto and Hamari [14] suggest that the effects of gamification could be smaller when using rigorous experimentations. In

sum, these reviews indicate there is no clear evidence of the effectiveness of gamified interventions. Nevertheless, no meta-analysis has been conducted yet. Quantifying the effect size of gamified interventions and identifying moderators of this effect would provide important information about the effectiveness of such interventions. Moreover, a meta-analysis appears as timely as there are now enough RCTs to conduct such analysis.

The present study

This study is the first to quantify the effect of gamified interventions on PA. Beyond the effect during or just after the intervention, we also seek to evaluate the long-term effect to determine the health relevance of these interventions. Indeed, we reasoned that to be considered effective, gamification must sustain its impacts over the long-term, and offer more than a short-term novelty effect [10]. The scalability of gamification is also a major issue because it would determine if gamification can be introduced in healthcare settings with patients or if it is more suited in prevention for healthy audiences. Finally, the mechanisms by which gamification works are not well understood. Yet, this step is critical to design efficient interventions.

The objectives of this systematic review and meta-analysis are to answer these research gaps by (1) evaluating the effect of gamified interventions on PA and SB, (2) assessing the long-term or sustained effect of gamified programs, (3) evaluating the scalability of gamification over different populations and (4) better understanding the mechanisms underlying the gamification effects.

METHODS

The present review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [16]. Following recommendations to minimize bias and provide evidence of a priori analyses intentions [17], the study was preregistered under the international prospective register of systematic reviews (PROSPERO; registration number: CRD42020186882) and on the Open Science Framework (OSF) [18]. Moreover, all materials and data are available on the OSF page of the project in order to facilitate reproducibility and transparency of this review [19].

Search strategy and information sources

In May 2020, we conducted a systematic literature search using five electronic databases – Pubmed, EMBASE, Scopus, Web of Science and the Cochrane Central Register of Controlled Trials (CENTRAL). We combined alternate terms and keywords representing synonyms for the outcomes (PA or SB), the intervention (gamification), the device (e-health), and the trial. The search strategy was reviewed by an academic librarian. All databases were searched individually, from 2010 (date of

the widespread adoption of the term “gamification” [9]) and the research was restricted to English-language texts. The complete search equations for all databases are available in the supplementary materials (Multimedia Appendix 1). In addition, we complemented our search with references harvesting from included studies and overview articles.

Eligibility criteria

Studies were eligible for inclusion if they were RCTs and if they met other criteria based on PICO (participants, interventions, comparators, and outcomes).

Participants. This review focused on the general population regardless of age, gender or health status (i.e., patients with chronic diseases were also included). We excluded studies involving participants with contra-indication to PA or disease preventing them from engaging in PA or understanding the principles of the game (intellectual and cognitive impairments).

Intervention. Of interest were digital interventions incorporating game elements and gamification techniques, such as points, levels, rewards, leaderboards, narratives, teams, etc. As introduced before, we clearly distinguished gamification and related constructs like serious games. For this reason, we excluded interventions based on active video games.

Comparators. Studies that attempted to compare gamified interventions to control groups without gamification elements in a randomized design were integrated in the review. These groups could be either inactive (non-exposed control group like a waiting list) or active (another intervention).

Outcomes. We included in the review studies assessing change in total PA, or leisure PA (quantity in MET.hour per week or MET.minute per week or in duration, energy expenditure (METs), moderate to vigorous physical activity (MVPA), steps count, walking time, active minutes) and change in times spent in SB (total time, during leisure time, work, in front of computer, in front of TV). These outcomes were continuous data either objectively measured (through accelerometers, pedometers, smartphones) or subjectively measured by self-reported questionnaires. Data measured objectively were always prioritized in the analyses over self-reported questionnaires, more susceptible to bias with a potential overestimation of PA [20].

In addition, studies were excluded if they came from a review, a commentary or a conference abstract, if they included data previously published in another study, if they were not randomized and controlled, if they were not written in English and if they were published before 2010.

Screening

Two authors (AM and AC) screened independently the titles and abstracts resulting from the search. Full texts of potential included studies were all checked before inclusion. Disagreement was resolved by discussion or by consulting a third author (MD), and agreement was measured with Kappa

statistic. A complete list of excluded studies is available on the OSF page of the project.

Data extraction

In the data collection process, AM and AC extracted data independently and blinded from each other using a pre-determined and tested template. Disagreement was resolved by discussion or by consulting a third author (MD). Extracted data included the results of each study and three types of potential moderators:

- Population-level moderators to assess the scalability of the intervention (population characteristics, age, gender, pathology)
- Intervention-level moderators to better understand gamification mechanisms (theoretical model used to develop the intervention, gamification elements, modality of the intervention (e.g., internet-based, smartphone app, presence of social incentives))
- Outcome-level moderators (outcomes, measure of PA or SB, device)

Risk of bias assessment

For each eligible study, two reviewers (AM and AC) assessed the risk of bias using the purpose-built Cochrane risk of bias tool (Table 8.5. in the Cochrane Handbook for Systematic Reviews of Interventions [21]), which evaluates seven domains (sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting and other potential risk of bias). A judgement of the potential risk of bias was made from the extracted information and rated as “high risk”, “low risk” or “unclear” if the related information was not available. These evaluations of bias are reported in the review and included in the analysis, and a measure of agreement with Kappa statistic was calculated. After the full assessment, we decided not to present the item “blinding of participants and personal” in the review because it was similar for all studies, which were rated as “high risk”, blinding being unfeasible for this kind of intervention.

Data synthesis and statistical analyses

First, data was synthesized in a qualitative review assessing the key elements of the studies and highlighting intervention differences. This qualitative review integrates all studies that met the eligibility criteria, including those for which we were unable to extract the data.

For the quantitative analysis, means and standard deviations of continuous PA or SB outcomes from individual studies were compiled when available or estimated using Hozo et al. [22] when median and interquartile range were reported. When the necessary data were not available in the original article, we first requested it to the authors. If data could still not be obtained, we extracted the data

from the graphs when available. If it was not possible, we excluded the study from the quantitative analysis.

A global meta-analysis was conducted to obtain a summary effect. In addition, when sufficient data was available (i.e., 4 studies or more reporting an outcome), we conducted different meta-analyses for each specific outcome (steps, MVPA, time in SB) and for the follow-up effect. To address the non-independence of data due to study effect, random-effects models [23] were preferred to the usual statistical tests. In addition, the Hartung-Knapp-Sidik-Jonkman method was used to reduce the production of false positive inherent of the DerSimonian-Laird method [24], and to obtain more robust estimates of variance. Continuous outcomes were analyzed using Standardized Mean Difference (SMD) to account for different measurement instruments or Mean Difference (MD) when the measurements were close enough. We computed Hedge's g [25] for effect sizes, that is similar to Cohen's d but corrects for small sample bias, which are recurrent in the studies included. In this way, a Hedge's g of 0.2 represents a small effect, 0.5 a moderate effect and 0.8 a large effect [26]. We computed SMDs both for outcome scores post-intervention (presented in the review) and change-from-baseline outcomes. Scores post-intervention effect sizes refer to treatment group results compared with control group results after interventions. Change-from-baseline scores effect sizes were calculated as a comparison between treatment group pre-post effect size and control group pre-post effect size.

For studies that included multiple outcomes, we kept in the main analysis the primary outcome targeted in the initial article. If none of the PA outcomes reported by a study were the primary ones, we selected the one that was the most relevant from the perspective of the intervention and the original experiment. In designs with multiple time measurements, the assessment that was the most proximal to the end of the intervention was conserved. A time assessment had to be realized more than two weeks after the end of the intervention to be included in the follow-up analysis.

Statistical heterogeneity was tested using forest plots and I^2 statistic, which is the most common metric for measuring the magnitude of between-study heterogeneity, and is easily interpretable (0% to 40%: might not be important; 30% to 60%: may represent moderate heterogeneity; 50% to 90%: may represent substantial heterogeneity; 75% to 100%: considerable heterogeneity) [27]. We conducted different influential analyses to address between-study heterogeneity. We first explored the presence of outliers, defined as studies with confidence interval that does not overlap with the confidence interval of the summary effect. We also performed Leave-one-out analyses which recalculate the summary effect several times with one study omitted each time. Finally, we performed a Baujat Plot [28] which is a diagnostic plot to detect sources of heterogeneity in the

meta-analysis by comparing the contribution of each trial in the pooled effect to the overall Cochran Q-test for heterogeneity.

We applied different methods to detect publication bias (Funnel plot, Egger's regression test [29], Duval & Tweedie's trim-and-fill procedure [30]). In addition, another approach to determine the evidential value of studies included in the analysis is to check the statistical power of the individual studies. To this end, we performed a sunset funnel plot [31] which is a funnel plot variant that visualizes the statistical power of each study included in the meta-analysis, based on the summary effect size.

Thus, sensitivity analyses were conducted to address studies with a high risk of bias, a strong heterogeneity in the sample or studies identified as outliers. Subgroup analyses were conducted to explore possible sources of heterogeneity and test for population differences. To this end, we conducted tests for subgroup differences using random-effects-model. In addition, moderation analyses were performed to explore the impact of potential explanatory variables and moderators on the effect size with meta-regressions when enough data were available (i.e., at least ten studies for each explanatory variable [21]). The results were expressed as regression coefficients estimates, 95% confidence intervals, and the *P* value.

For cross-over trials, we first checked in the original texts of studies that neither carry-over nor period effects were problematic. For cluster-randomized trials, we checked if the influence of the different clusters was not too important, analyzing values of intra-class correlation coefficient in studies. Then, in the absence of enough information in the published articles, we addressed these studies as traditional parallel trials. Nevertheless, this procedure increases the probability of unit-of-analysis error. This is why we performed sensitivity analyses without clusters and cross-over trials to test the robustness of our results.

A summary figure of the analytic procedure is available in the supplementary materials (Fig. S1 in Multimedia Appendix 1). Analyses were performed on R (The R Project for Statistical Computing) with the *dmetar* package [32]. Risk of bias summary and risk of bias graph were made via the *robvis* R package [33].

RESULTS

Characteristics of included studies

The authors screened titles and abstracts of 1480 results and 51 full-text articles were assessed for eligibility according to the inclusion criteria. Finally, 21 articles [32-52] were included in the qualitative analysis and 16 in the meta-analysis (Fig. 1). The kappa measure of agreement for the screening phase was of 0.64 between the two authors, reflecting good agreement [55].

Table 1 gives an overview of included studies characteristics. The sixteen studies included in the quantitative analysis involved 2407 participants aged from 9 to 73 years old ($M = 35.7$ years) with sample sizes ranging from 20 to 602. Overall, fifteen studies (71%) included adults' participants and six (29%) included adolescents (less than eighteen-year-old). Five (24%) studies included patients with chronic diseases (i.e., obesity, type 2 diabetes, rheumatoid arthritis, and cardiovascular disease).

Table 1 Characteristics of included studies

Study	Participants	Intervention	Theory	PA outcomes
Allam et al. (2015)	Rheumatoid arthritis patients, N = 157 (<i>Mage</i> = 59.1 years, <i>SD</i> = 9.44, 48% of male participants)	The intervention was a website with the aim to provide specific knowledge on rheumatoid arthritis to patients. Gamification was used to motivate the patients to use the platform. Duration = 8 weeks	-	Self-reported PA (Exercise Behaviors Scale)
Corepal et al. (2019) *	Adolescents, N = 224 (aged 12 – 14 years, 47% of male participants)	The <i>StepSmart Challenge</i> was a web-based intervention that used gamification strategies to encourage and support PA behavior change. Duration = 22 weeks	SDT	Daily step count and MVPA (min/day) objectively measured (Actigraph GT3x)
Dadaczynski et al. (2017) *	Adults workers in an automobile manufacture, N = 144 (65% of male participants)	<i>Healingo Fit</i> had the objective to promote low levels of PA, using a tracking-based approach measuring PA with a Fitbit pedometer and a gamified intervention accessible by desktop and mobile devices. Duration = 6 weeks	SCT, HAPA	TPB, Self-reported VPA, MPA, Minutes walked (min/week) (IPAQ)
Direito et al. (2015) *	Adolescents, N = 35 (<i>Mage</i> = 15.7 years, <i>SD</i> = 1.2, 45% of male participants, BMI = 22.85)	<i>Zombies, Run! 5K Training app</i> was a fully automated training program designed to improve fitness, combined with an immersing and fun story. Duration = 8 weeks	SDT	MVPA, VPA, MPA, LPA and sedentary time (min/day) objectively measured (Actigraph GT3x), self-reported PA (PAQ-A)
Edney et al. (2020) *	Adults, N = 284 (<i>Mage</i> = 41.2 years, <i>SD</i> = 11.2, 25% of male participants, BMI = 30.1)	<i>Active Team</i> was a mobile app designed to encourage inactive adults to meet PA guidelines. Gamification and social features were implemented to increase the social comparison, support, and influence between participants. Duration = 12 weeks	SCT	MVPA (min/day) objectively measured (GENEActiv) and self-reported PA (Active Australia Survey)
Garde et al. (2015)	Adolescents, N = 47 (<i>Mage</i> = 10.3 years, <i>SD</i> = 1.9, 34% of male participants, BMI z-score = 0.35)	<i>MobileKids Monster Manor</i> was a mobile exergame synchronized with an external activity monitor. The overall goal was to complete the story with physical activity and steps. Duration = 1 week	SDT	Daily step count and active minutes/day objectively measured (Tractivity)
Garde et al. (2016) *	Adolescents, N = 56 (<i>Mage</i> = 11.3 years, <i>SD</i> = 1.2, 62% of male participants, BMI z-score = 0.28)	<i>MobileKids Monster Manor</i> was a mobile exergame synchronized with an external activity monitor. The overall goal was to complete the story with physical activity and steps. Duration = 1 week	SDT	Daily step count and active minutes/day objectively measured (Tractivity)

Garde et al. (2018)	Adolescents, N = 37 (<i>Age</i> = 10.6 years, 43% of male participants, BMI z-score = 0.21)	<i>MobileKids Monster Manor</i> was a mobile exergame synchronized with an external activity monitor. The overall goal was to complete the story with physical activity and steps. Duration = 2 weeks	SDT	Daily step count and active minutes/day objectively measured (Tractivity)
Gotsis et al. (2013)	Adults, N = 142 (<i>Age</i> = 35.7 years, <i>SD</i> = 9.5, 33% of male participants, BMI = 28)	<i>Wellness Partners</i> was a web-based intervention in the form of a physical activity diary with social networking and game features. Duration = 10 weeks	-	Self-reported PA (days/week, 1 item)
Gremaud et al. (2018) *	Adult office workers, N = 144 (<i>Age</i> = 40.5 years, <i>SD</i> = 11.4, 76% of male participants, BMI = 29.7)	<i>MapTrek</i> was a mobile health platform that gamified Fitbit use for the purpose of promoting physical activity by placing users in a series of virtual walking races. Duration = 10 weeks	SDT, SCT	Daily step count and daily active minutes count objectively measured (Fitbit Zip activity monitor)
Haque et al. (2020)	Adult office workers, N = 84 (aged 24-49 years, 54% of male participants, BMI = 24.8)	The objective of the <i>iGO</i> mHealth app was to encourage employees to walk more often and to break up long sitting periods during working hours by the introduction of virtual rewards. Duration = 4 weeks	SDT	Self-reported amelioration of PA (1 item)
Höchsmann et al. (2019) *	T2DM and obese patients, N = 35 (<i>Age</i> = 58.5 years, 53% of male participants, BMI = 32)	The intervention was a mobile app including a storyline, virtual rewards, individualized exercises, and daily PA promotion through a game. Duration = 24 weeks	Taxonomy of behavior change techniques	Daily step count objectively measured (Garmin Vivofit 2)
Kurtzman et al. (2018) *	Obese adults, N = 196 (<i>Age</i> = 41.4 years, <i>SD</i> = 12.2, 13% of male participants, BMI = 36.2)	Participants were in teams of two and had to complete weekly goal targets to win points and badges. Duration = 24 weeks	Behavioral economics	Mean step count objectively measured (Withings wrist-worn device)
Leinonen et al. (2017) *	Adolescents, N = 496 (<i>Age</i> = 17.8 years, <i>SD</i> = 0.6, 100% of male participants, BMI = 23.1)	The intervention was an app proposing a mixed-reality conquering game based in which physical and social activities are rewarded. Duration = 24 weeks	Transtheoretical model	Daily MVPA and daily sedentary time objectively measured (Polar Active)
Maher et al. (2015) *	Adults, N = 110 (<i>Age</i> = 35.6 years, <i>SD</i> = 12.4, 42% of male participants)	<i>Active Team</i> was a mobile app designed to encourage inactive adults to meet PA guidelines. Gamification and social features were implemented to increase the social comparison, support, and influence between participants. Duration = 8 weeks	TPB	Self-reported MVPA, VPA, MPA and minutes walked (min/week, Active Australia Survey)

Nishiwaki et al. (2014) *	Adults, N = 20 (<i>Mage</i> = 31 years, <i>SD</i> = 3, 30% of male participants, BMI = 21.5)	Participants wore an activity monitor with computerized game functions like a story, a character and objectives. Duration = 6 weeks	-	Daily step count and MVPA (METs h/day) objectively measured (Lifecorder EX)
Patel et al. (2017) *	Adults, N = 200 (<i>Mage</i> = 55.9 years, <i>SD</i> = 9.9, 44% of male participants, BMI = 27.1)	Participants were entered into a game with their family in teams and had to complete weekly goal targets to win points and badges. Duration = 12 weeks	Behavioral economics	Daily step count objectively measured (Withings wrist-worn device)
Patel et al. (2019) *	Overweight and obese adults, N = 602, (<i>Mage</i> = 38.7 years, <i>SD</i> = 10.4, 69% of male participants, BMI = 29.6)	Participants had to complete weekly goal targets to win points and levels. There were three versions of the intervention: support, collaboration, and competition. Duration = 24 weeks	Behavioral economics	Daily step count objectively measured (Withings wrist-worn device)
Paul et al. (2016) *	Stroke survivors' patients, N = 23, (<i>Mage</i> = 55.8 years, <i>SD</i> = 10.7, 48% of male participants, BMI = 24.5)	In the <i>STARFISH</i> app, participants had to complete their PA objectives to improve their avatar. Duration = 6 weeks	Control theory, Michie's Taxonomy of behavior change	Daily step count, sedentary time, walking time (min/week) objectively measured (ActivPAL)
Thorsteinsen et al. (2014) *	Adults, N = 21, (<i>Mage</i> = 55.3 years, <i>SD</i> = 11.2, 52% of male participants)	The intervention <i>Lifestyle Tool</i> consisted of a rule-based website designed to help people plan and monitor their physical activity. The tool incorporated social and individual gaming components to increase motivation and engagement. Duration = 12 weeks	-	Self-reported weekly activity minutes (daily report form)
Zuckerman & Gal-Oz (2014) *	Students, N = 59, (<i>Mage</i> = 23.4 years, <i>SD</i> = 1.4, 25% of male participants)	<i>StepByStep</i> was an accelerometer-based mobile application with virtual rewards and social comparison intended to motivate people to incorporate more walking into their daily routine. Duration = 1.5 week	SDT	Walking time (min/day) objectively measured (smartphone accelerometer)

Asterisks indicate studies included in the meta-analysis. MVPA moderate-to-vigorous physical activity, VPA vigorous physical activity, MPA moderate physical activity, LPA light physical activity, PA physical activity, SDT self-determination theory, SCT socio-cognitive theory, TPB theory of planned behavior, HAPA health action process approach, T2DM type 2 diabetes mellitus

Eight trials were conducted in Europe, five in the USA, three in Australia and New-Zealand, three in Canada and two in Asia. Studies were published from 2013 to 2020 with eight (38%) after 2018. Most studies were either based on internet (n = 5), or on a smartphone app (n = 11), or both (n = 4). The study of Nishiwaki and colleagues [41] had the particularity to use a pedometer with computerized game functions. The duration of the intervention varied from one to twenty-four weeks with a mean of 11.8 weeks, and the most common length was 24 weeks. The most used game mechanics were virtual rewards like badges, medals or trophies (n = 16; 76%), teams or leagues (n = 13; 62%), a way to visualize progress like feedbacks or game statistics (n = 10; 48%) and points or scores (n = 10; 48%). Almost all studies included social incentives such as team collaboration, social

networking, and messaging facilities, in their intervention (86%).

In the majority of studies, interventions were based on theoretical models. Six studies were based on socio-cognitive models (i.e. the transtheoretical model [56], the social cognitive theory [57], the theory of planned behavior [58], the Health Action Process Approach (HAPA) [59]), five on the self-determination theory [60] and three on behavioral economics models.

Outcomes measured in trials were diverse: they used either total PA duration and/or MVPA duration, SB duration, daily steps count, walking duration or active minutes count. Thirteen experiments measured PA objectively using devices such as tri-axial accelerometers ($n = 7$), wearables devices for general population (i.e., Fitbit, Garmin, Polar, and Withings monitors; $n = 6$), pedometers ($n = 1$), or smartphones ($n = 1$), and six assessed PA with self-reported measures (IPAQ, PAQ-A, Active Australian Survey, Exercise Behavior Scale or other). Six trials completed a follow-up assessment from 12 to 30 weeks ($M = 14.4$ weeks) after the end of the intervention.

Five studies were excluded from the meta-analysis and were only integrated in the qualitative review. We were unable to extract the results of two studies [50,52]. Studies of Gotsis et al. [53] and Haque et al. [36] did not assess PA with an accepted and validated metric and could not be integrated in the meta-analysis. Lastly, the paper of Allam et al. [34] was excluded because it was the care that was gamified and not PA.

Fig. 1 PRISMA flow chart of the literature search and screening process

Risk of bias

The two authors extracted the risk of bias data with a kappa coefficient of 0.79, synonym of excellent agreement [55]. Tables 2 and 3 give authors' judgements about each risk of bias item presented as percentages across all included studies in the meta-analysis, and an overview of the different bias for each study. One study [44] was rated as high risk for the sequence generation because assignments were based on recruitment order. For this reason, this study was at high risk for allocation concealment too. Four studies [38,43,44] were at high risk of bias for the 'blinding of outcome assessment' item because they measured PA using only self-reports. Six studies [33,37,43,44,49] were at high risk of bias for the 'incomplete outcome data' item because they reported high drop-out and did not included intention-to-threat analyses. Six studies [39,42,43,45,49] were rated at unclear risk for the 'selective outcome reporting' item because they have not been pre-registered or published in a protocol-study. Lastly, two studies had other high risks of bias. The first one [41] was a cross-over trial conducted without washout condition, and the other one [35] was a cluster randomized trial with no control of clustering, no consideration of the clustering in the statistical analysis and no test of baseline differences between groups.

Table 2 Risk of bias summary for studies included in the meta-analysis

	Risk of bias						Overall
	D1	D2	D3	D4	D5	D6	
Corepal et al., 2019	+	+	+	X	+	X	X
Dadaczynski et al., 2017	+	-	X	X	+	+	X
Direito et al., 2015	+	+	+	+	+	+	+
Edney et al., 2020	+	+	+	+	+	+	+
Garde et al., 2016	+	+	+	X	-	+	X
Gremaud et al., 2018	+	+	+	+	+	+	+
Höchsmann et al., 2019	+	+	+	+	+	+	+
Kurtzman et al., 2018	+	+	+	+	+	+	+
Leinonen et al., 2017	+	+	+	X	+	+	X
Maher et al., 2015	+	+	X	+	+	+	X
Nishiwaki et al., 2014	-	-	+	+	-	X	X
Patel et al., 2017	+	+	+	+	+	+	+
Patel et al., 2019	+	+	+	+	+	+	+
Paul et al., 2016	X	X	+	+	-	+	X
Thorsteinsen et al., 2014	+	-	X	X	-	+	X
Zuckerman & Gal-Oz, 2014	-	-	+	+	-	+	+

D1: Random sequence generation
 D2: Allocation concealment
 D3: Blinding of outcome assessment
 D4: Incomplete outcome data
 D5: Selective reporting
 D6: Other bias

Judgement
 X High
 - Unclear
 + Low

Table 3 Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies

Summary effect

Overall, the SMD post-intervention for all PA outcomes (MVPA, daily step count, number of active minutes, and walking time) was of $g = 0.43$ (95%CI [0.03, 0.82], $I^2 = 86\%$), (Fig. 2) representing a

statistically significant small to medium effect. In the same vein, we found a statistically significant SMD effect of $g = 0.38$ (95%CI [0.07, 0.69], $I^2 = 79\%$) for pre-post change scores.

Only three studies [39,44,48] assessed sedentary time as an outcome. Due to this small sample, the meta-analysis was not performed on this outcome.

Fig. 2 Forest plot for the effect of gamification versus control on post-intervention PA outcomes (moderate-to-vigorous physical activity (MVPA), daily step count, number of active minutes, and walking time). *SD* standard deviation, *Std* standardized, *IV* inverse variance, *CI* confidence interval, τ^2 , χ^2 , and I^2 measures of between-study heterogeneity

Outliers and influential analyses

In the first results, a substantial statistical heterogeneity was observed. To address between-study heterogeneity, we first looked for the presence of outliers. Three studies were considered as outliers [35,37,41] and after removing them we still obtained a significant effect of $g = 0.34$, 95%CI [0.17, 0.51] with a moderate heterogeneity ($I^2 = 58\%$).

Leave-one-out analyses showed that sequential removal of each study did not had an important impact on the general effect size with effect sizes ranged from $g = 0.33$ (95%CI [0.00, 0.66], $I^2 = 84\%$) to $g = 0.48$ (95%CI [0.13, 0.83], $I^2 = 78\%$) (Fig. S2 in Multimedia Appendix 1).

The Baujat plot (Fig. S3 in Multimedia Appendix 1) indicated that four studies explained more heterogeneity than the others, and more especially the study of Corepal et al. [35] with a heterogeneity contribution of 40.13 and an effect size influence of 3.27.

We therefore excluded studies with a high or unclear risk of bias. After doing so, the effect was not significant ($g = 0.33$ (95%CI [-0.16, 0.81], $I^2 = 78\%$).

As introduced before, the inclusion of cross-over and cluster randomized trials in the meta-analysis may lead to biases. Thus, we excluded these designs from the analysis and obtained a statistically significant effect of $g = 0.49$ (95%CI [0.05, 0.92], $I^2 = 67\%$).

Finally, we decided to exclude articles of Corepal et al. [35] and Nishiwaki et al. [41] in sensitivity analyses (which will be repeated afterwards for each analysis) considering their influence on the pooled result, their contribution on overall heterogeneity and their huge risk of bias (no control of clustering, no statistical consideration of clustering and no test of baseline differences between groups for Corepal et al., and no washout period and very low power for Nishiwaki et al.). After doing so, we obtained a statistically significant effect of $g = 0.42$ (95%CI [0.14, 0.69], $I^2 = 74\%$).

Subgroup analyses

We found no statistical differences in the effects between studies with subjects with chronic diseases or healthy subjects ($Q = 0.73$, $P = 0.39$), between adults and adolescents ($Q = 0.26$, $P = 0.61$), between studies with objective (devices) or self-reported PA outcomes ($Q = 0.23$, $P = 0.63$), between active or inactive control groups ($Q = 0.01$, $P = 0.92$), and between short or long-term interventions (less or more than 12 weeks) ($Q = 0.60$, $P = 0.44$).

When performing the sensitivity analysis, there was a statistically significant effect of intervention on PA for adults ($g = 0.36$, 95%CI [0.03, 0.69], $I^2 = 71\%$), on healthy people ($g = 0.35$, 95%CI [0.15, 0.55], $I^2 = 63\%$), when the PA measure was objective ($g = 0.45$, 95%CI [0.08, 0.82], $I^2 = 80\%$), self-reported ($g = 0.35$, 95%CI [0.21, 0.50], $I^2 = 0\%$), when the control group was active ($g = 0.32$, 95%CI [0.05, 0.59], $I^2 = 49\%$), and for short interventions of 12 weeks (equivalent of a three months program) or less ($g = 0.31$, 95%CI [0.10, 0.52], $I^2 = 45\%$).

Meta-regressions

Neither the age of participants ($\beta = 0.01$, 95%CI [-0.02, 0.04], $P = 0.39$), their gender ($\beta = 0.01$, 95%CI [-0.01, 0.02], $P = 0.47$), their BMI ($\beta = -0.04$, 95%CI [-0.16, 0.09], $P = 0.53$), the duration of the intervention ($\beta = -0.01$, 95%CI [-0.06, 0.04], $P = 0.74$), nor the number of game mechanics included in the intervention ($\beta = 0.01$, 95%CI [-0.17, 0.19], $P = 0.91$) were statistically significantly associated with increase of PA. Lack of data precluded further meta-regressions, such as comparisons on leisure physical activity, or test of moderators, like the impact of social incentives or the theoretical model used to develop the intervention.

Publication bias

First, an inspection of the funnel plot shows that the individual studies' effect sizes were relatively symmetrically distributed around the pooled effect size. This observation is supported by the Egger's test of the intercept which indicates no asymmetry in the funnel plot ($b0 = 1.38$, 95%CI [-0.83, 4.77], $P = 0.19$). We then applied a bias-correction technique, the trim and fill method, which indicated that three studies were missing at the bottom left of the funnel plot to obtain a full symmetry. After

imputing the effect sizes corresponding of these missing studies to obtain a totally symmetrical funnel plot (Fig. 3), the bias corrected summary effect was of $g = 0.24$ (95%CI [-0.24, 0.73]).

Fig. 3 Funnel plot after Trim and Fill bias correction. A filled circle represents an included study, and an empty circle represents a “missing” study.

Finally, the sunset funnel plot (Fig. 4) showed huge differences in power between studies, with some characterized by very low statistical power (seven under 45% power and four under 18%). The median power of all tests was of 63%.

Fig. 4 Power-enhanced funnel plot. White circles represent included studies. δ true effect size, $\text{med}_{\text{power}}$ the median power of all tests, $d_{33\%}$ effect size needed for achieving 33% of median power, $d_{66\%}$ effect size needed for achieving 66% of median power, E expected number of positive studies, O observed number of positive studies, p_{TES} test of excess significance P -value

Secondary analyses

Follow up. There was first no statistically significant effect of gamified interventions on total physical activity (MVPA and daily step count) after follow-up periods with a SMD of $g = 0.09$ (95%CI [-0.07, 0.26], $I^2 = 21\%$). When we performed the sensitivity analysis, gamification significantly increased PA (MVPA and daily step count) in follow-up (from 12 to 24 weeks after the end of the intervention) ($g = 0.15$ (95%CI [0.07, 0.23], $I^2 = 0\%$) (Fig. 5).

Fig. 5 Forest plot for the effect of gamification versus control on PA outcomes (moderate-to-vigorous physical activity (MVPA), and daily step count) after a follow-up period (from 12 to 24 weeks after the end of the intervention). *SD* standard deviation, *Std* standardized, *IV* inverse variance, *CI* confidence interval, τ^2 , χ^2 , and I^2 measures of between-study heterogeneity

Steps. First, we found no statistically significant effect of gamified interventions on steps outcomes with a SMD of $g = 0.53$ (95%CI [-0.09, 1.15], $I^2 = 89\%$), but a significant improvement of the number of daily steps with a MD of +1420.57 steps per day (95%CI [435.41, 2405.73], $I^2 = 95\%$). When excluding the two studies in the sensitivity analysis, we obtained a statistically significant effect of gamification on daily steps of $g = 0.49$ (95%CI [0.05, 0.93], $I^2 = 75\%$) (Fig. 6) and a statistically significant MD of +1609.56 steps per day (95%CI [372.39, 2846.73], $I^2 = 86\%$) (Fig. 7).

Fig. 6 Forest plot for the effect of gamification versus control on steps outcomes (daily step count and walking time). *SD* standard deviation, *Std* standardized, *IV* inverse variance, *CI* confidence interval, τ^2 , χ^2 , and I^2 measures of between-study heterogeneity

Fig. 7 Forest plot for the mean difference of daily steps between gamification and control. *SD* standard deviation, *IV* inverse variance, *CI* confidence interval, τ^2 , χ^2 , and I^2 measures of between-study heterogeneity

MVPA. There was no statistically significant effect of gamification on MVPA with a SMD of $g = 0.09$ (95%CI [-0.57, 0.74], $I^2 = 93\%$). We still obtained a no statistically significant effect of $g = 0.31$ (95%CI [-0.19, 0.80], $I^2 = 82\%$) in the sensitivity analysis (Fig. 8).

Fig. 8 Forest plot for the effect of gamification versus control on moderate-to-vigorous physical activity (MVPA). *SD* standard deviation, *Std* standardized, *IV* inverse variance, *CI* confidence interval, τ^2 , χ^2 , and I^2 measures of between-study heterogeneity

Summary of findings (GRADE)

The quality of evidence (GRADE) in the included studies for short term PA, long-term PA, MVPA, steps and daily steps was scored from high to low (Table 4). The quality was downgraded for some outcome because of high heterogeneity, high risk of bias or imprecision due to large confidence intervals. Summaries of the various meta-analysis conducted in this review both on post-intervention scores and pre/post intervention change scores are presented in the supplementary materials (Fig. S4, Fig. S5, Fig. S6 in Multimedia Appendix 1).

Table 4 Summary of findings

Outcome	Number of participants (number of studies)	SMD or MD [95% CI]	Quality of evidence (GRADE)
---------	---	-----------------------	--------------------------------

General PA	2197 (14)	0.42, [0.14, 0.69]	Low ^{1,2,3}
Long-term PA (follow-up)	1306 (5)	0.15, [0.07, 0.23]	High
MVPA	739 (4)	0.31, [-0.19, 0.80]	Low ^{1,2,3}
Steps	1438 (9)	0.49, [0.05, 0.93]	Low ^{1,2,3}
Daily steps	1235 (7)	1609.56, [372.39, 2846.73]	Moderate ^{1,3}

GRADE: Grading of Recommendations Assessment, Development, and Evaluation

¹ Downgraded because of high heterogeneity

² Downgraded because of risks of bias

³ Downgraded because of imprecision (large confidence intervals)

DISCUSSION

Main findings

Summary effect. This meta-analysis of randomized controlled trials, including 16 studies and 2407 participants, revealed a statistically significant effect of gamified interventions on average of 12 weeks on total PA ($g = 0.42$ (95%CI [0.14, 0.69] after sensitivity analyses). This effect was small to medium, suggesting effectiveness of gamified interventions in promoting PA in both healthy and subjects with chronic diseases. This significant effect is robust as it persisted even after the different influence analyses were performed. Moreover, the effect was statistically significant both for objective measures of PA ($g = 0.45$, 95%CI [0.08, 0.82]) and self-reported ones ($g = 0.24$, 95%CI [0.08, 0.39]) after sensitivity analyses. These results are important considering the assets of gamification, which has the advantages (a) to reorganize existing activity rather than adding additional demands to people's lives [12], (b) to be easily implemented in natural contexts, and (c) to address a consequent number of participants through one unique program.

Long-term effect. When we analyzed the long-term effect of these interventions based on the follow-up measures of PA, carried out from 12 to 24 weeks ($M = 14.4$ weeks) after the end of the intervention, we found a statistically significant very-small to small effect size of $g = 0.15$ (95%CI [0.07, 0.23]) after sensitivity analyses. These results indicate that the effect of gamification persists after the end of the program, suggesting that it is not just a novelty effect due to the playful nature of gamification. However, this long-term effect is volatile, decreasing with time after the end of the intervention.

Scalability of gamified interventions. The absence of subgroup differences or effect of age, gender, and BMI on the pooled effect suggests a good scalability of gamified interventions, which can be used for several types of publics. Thus, gamification may not only be efficient in young healthy individuals but can target any kind of population regardless of their age or health status.

In sum, gamified interventions appear as an efficient tool to improve the PA of various population

but with a moderate sustainability of the effect post-intervention. Nevertheless, if many PA interventions increase PA levels in the short term, translating these temporary changes into long-term physical activity participation continues to be a challenge for PA research [61]. With that in mind, the potential of gamification for PA increase in the long-term, even minimal, is particularly important and promising in the area of PA interventions. In addition, the cost-effectiveness ratio of gamified interventions may be better than many current interventions, considering the ease of implementation and generalizability of gamification.

Additional findings

Effect of gamification on the step count. If the overall effect of gamified interventions on PA is positive, they increase the step count more than MVPA. Indeed, after sensitivity analyses, the meta-analyses revealed on the one hand a statistically significant effect ($g = 0.49$ (95%CI [0.05, 0.93]) of gamification for steps outcomes with a statistically significant improvement of 1609.56 steps per day (95%CI [372.39, 2846.73]) for participants benefiting from gamified intervention versus control group. On the other hand, no statistically significant effect of gamified interventions on MVPA was found ($g = 0.31$ (95%CI [-0.19, 0.80])). This can be explained by the game metrics and mechanics of the interventions included in the review that are mainly focused on the step count of participants. Few interventions were directly targeted MVPA or prolonged sedentary times. In the included studies, only two interventions [37,48] integrated multi-PA intensity goals and mechanics, notably with physical exercises or running sessions in the game, and just one introduced sedentary breaks as a gaming part [36]. In other words, participants played most of the time with their number of steps and had generally to walk more to make points and play the game. This results in an increase of the walking time but not necessarily in more intense PA.

These findings are interesting considering potential health benefits of increasing the number of daily steps by 1600 thanks to gamification. Indeed, previous work showed that walking was statistically associated with declines in all-cause and cardiovascular mortality [62–65], and an improvement in body composition [66]. Moreover, Oja et al. [64] suggested that any walking exposure is beneficial for cardiovascular health, endorsing the idea that the most important is more global PA regardless of the intensity [7,67] even when this activity only includes walking [62]. In comparison, a previous study evaluating the effectiveness of activity trackers with and without incentives to increase physical activity [68] showed a significant improvement of 1050 daily steps for the cash incentives intervention versus control (95%CI [600, 1490]) but no statistically difference for the Fitbit only group (340 daily steps, 95%CI [-100, 790]). In the light of these results, gamified interventions appear as an added value compared to current interventions. Especially considering that 40% of the

volunteers of this study abandoned their Fitbit monitor within 6 months, gamification is in addition a way to keep participants involved and motivated within the intervention.

Duration of intervention. Our meta-regression analysis did not find association between the observed effect of gamification on PA and intervention length. However, while no statistical significant effect of gamification for intervention lengths superior to 12 weeks was found ($g = 0.66$, 95%CI [-0.69, 2.01]), the meta-analysis revealed a statistically significant effect of gamified interventions of 12 weeks or less on global PA ($g = 0.31$, 95%CI [0.10, 0.52]). According with a previous meta-analysis that reported significant positive effects of smartphone apps on PA only when used over a short-term period of less than three months [69], these results suggest that a condensed intervention could benefit more than a longer one, which could become redundant, boring and exhausting for participants in the long run.

Statistical Heterogeneity. The meta-analysis also revealed considerable statistical heterogeneity. This heterogeneity may be explained by differences in study quality, diversity of designs and variations in study populations. Indeed, even if we did not find statistically differences in subgroup analyses, some analyses were performed with a small number of studies. We cannot rule out the potential variations of effect between adults and adolescents, disease-free participants or patients. In addition, these subgroups and characteristics may not explain all the variance of the interventions. Indeed, demographic data often do not fully explain differences in the effectiveness of interventions [70,71] and more precise socio-psychological variables such as personality traits or motivational factors could explain this poorly understood significant heterogeneity. Moreover, the risk of bias analysis and the sunset funnel plot showed substantial differences in the qualities of included studies that can influence the heterogeneity. Finally, various trial designs were used (i.e., parallel RCT, Cluster RCT, Cross-over RCT) that can contribute too to overall statistical heterogeneity.

Better understanding gamification mechanisms

This meta-analysis is informative regarding the effectiveness of gamified interventions. In view of the observed heterogeneity, the next step will be to investigate its causes from an interventional and theoretical perspective. Gamified interventions involve multiple interacting elements, and it is crucial to estimate the weight of each element in the behavior change process, and how they interact with each other. Is it game mechanics, the implementation of behavior change techniques, or the presence of social interactions that make gamification effective? Unfortunately, the small amount of studies included in the meta-analysis impeded us from conducting in-depth moderation analyses to answer this question. To better understand these relations, it is essential that both the development and the assessment of interventions be central, transparent, evidence-based, context-aware, and research-

oriented [72]. Moreover, if theoretical psychological models are often mentioned in the introduction of articles included in the review (see table 1), few have investigated the psychological mechanisms of their interventions on the field [36,46,48,73]. Future studies should explicitly discuss motivational theory and systematically test the effect of gamification on psychological outcomes known to be involved in behavior change (e.g., self-efficacy, attitudes, intention), in order to better understand its mechanisms. As introduced before, the consideration of personality traits and psychological variables to determine behavioral phenotypes [70] are a promising way to evaluate participant's responses to interventions.

Another explanation of the high heterogeneity between studies could be the role of game mechanics and social incentives in the intervention. We observed a variety of game mechanics within the intervention panel and almost all introduced social features. However, we postulate that beyond these constructs, the need to individualize game processes is vital. Joining Nicholson's vision of meaningful gamification [74] fighting the common belief that the integration of game mechanics is a self-fulfilling process to engage participants, we suggest that, to be efficient, game strategies need to focus on specific psychological variables which can differ according to the population characteristics such as age, socio-economic or health status. Thus, the interventions that are most likely to succeed would be those that consider the characteristics of the target population and try to create games that make sense for participants. The study of Höchsmann et al. [37] is an excellent example. They obtained the highest effect size of the sensitivity analysis ($g = 2.41$ (95%CI [1.52, 3.31]) with an intervention specifically designed for type 2 diabetes patients. They considered the characteristics of the patients and the age group needs in offering tailored content and a game world adapted to their attributes and expectations. To some extent, the situation would be similar for the effect of social interactions within the games. It is not enough to set up social incentives in digital interventions to capitalize on social benefits such as social support and social influence [75]. We suggest that the social connection between participants is primordial and will influence social incentives during the game.

Other perspectives and recommendations for future research

The findings from this meta-analysis allow us to draw and discuss future work concerning gamification of PA and SB. The long-term effect is nowadays the main challenge of health interventions. It is thus essential to investigate the evolution of the effect of gamified interventions over time. To this end, there is a need of more long-term follow-up measurements. In addition, the potentialities of digital technologies and their capacity to collect a large amount of real-world data

could be used to assess the evolution patterns of the effect, allowing to identify in detail its sustainability, evolvement, or even make forecasts via time series analyses. Second, there is an urgent need to develop gamified interventions that target SB. Gamification has clearly the potential to address this issue and future research should implement specific game mechanics affording participants to take more sedentary breaks. Third, the development of gamified interventions may be based on tactics that have been scientifically proven to positively influence health outcomes. Finally, future trials should be conducted with more adequately powered sample sizes and be strictly multiple arms randomized controlled trials to isolate the effects of gamification elements.

Implications for practice

In the light of our results, gamified interventions appear to be a promising avenue to promote PA in different publics, both in prevention for healthy people and in treatment of chronic diseases. Regarding subjects with chronic diseases, gamified interventions have many benefits like (a) empowerment of participants by improving their self-management skills, (b) an effect across broad audiences enabling to target different types of pathologies, and (c) an everyday life fit and easy implementation. Again, like other digital health processes, gamification makes it possible to address more patients and more especially those who are isolated from healthcare facilities. One important element seems the need to emphasize the population target and adapt the content according to participants age, the pathology, and the associated constraints.

Nevertheless, in view of the volatile post-intervention effect, the current article suggests that a one-shot intervention is not sufficient. A more interesting design would be to address multiple “gamification doses” during and/or after the course of treatment to obtain a sustainable implementation of the PA behavior. This configuration would also provide an ideal duration of intervention to avoid exhausting the participants.

More generally, we suggest that gamified interventions should be a part of a larger intervention comprising supervised PA and health education to obtain a sustainable behavior change. Here, gamification would be a process permitting to increase overall PA (structured PA and daily-life PA) and potentially developing an intrinsic motivation towards PA [73]. Like the intervention proposed by Allam et al. [34], health education could be gamified too by the introduction of points and quizzes to improve patients adherence.

Strengths and limitations

To our knowledge, this is the first meta-analysis to quantitatively evaluate the effect of gamification on PA. This review has other several strengths. First, we conducted a comprehensive search strategy using multiple databases in collaboration with an academic librarian. Then, all stages of the review

have been realized independently by two reviewers (screening and data extraction). Lastly, various and novel publication bias analyses and influence analyses were conducted in parallel of different subgroups analyses and meta-regressions.

However, some limitations of the current work have to be mentioned. Overall, the meta-analysis included a small number of studies, and some articles were feasibility or pilot trials. That is why several trials included small sample sizes and were highly underpowered. Some studies were conducted with a high risk of bias. Finally, we included in the main analysis diverse PA outcomes evaluating similar constructs, but which are slightly different in practice. Moreover, not all included outcomes were objectively measured. As the field matures and new trials are published, an update of this work will be important to confirm these preliminary results.

CONCLUSION

To conclude, gamified interventions appear to be a promising avenue to promote PA to various publics. Influencing primarily the number of daily steps of the participants, gamification is an interesting way to improve daily PA. However, if the effect of gamification persists in follow-up, this effect decreases with time with a more volatile long-term effect. The integration of gamification in more global healthcare interventions could be a way to address this limited sustainability. Future rigorous trials are needed to explore these perspectives.

Acknowledgements

The authors wanted to thank the librarian of the Clermont Auvergne University for the assistance for developing the search equations.

Funding

The work of AM is supported by an ANRT grant (Cifre PhD Thesis) and by the company Kiplin. None of the sponsors were involved in any other aspect of the project, such as the design of the project's protocol and analysis plan, the collection and analyses. The funders had no input on the interpretation or publication of the study results. The authors wanted also to thank the challenge 3 I-SITE Clermont Auvergne Project 20-25 for their grant.

Conflicts of interest

The authors declare that they have no conflicts of interest relevant to the content of this review.

Availability of code, data and materials

All data, materials and R code are available on the OSF page of the project [18].

Authors' contributions

AM designed the study protocol and search strategy in consultation with other authors. AM and AC conducted the screening, data extraction and risk of bias assessment. AM and BP developed the meta-analytic procedure and AM realized the statistical analysis. The first draft of the manuscript

was written by AM and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

REFERENCES

1. P.C. Hallal, L.B. Andersen, F.C. Bull, R. Guthold, W. Haskell, U. Ekelund, J.R. Alkandari, A.E. Bauman, S.N. Blair, R.C. Brownson, C.L. Craig, S. Goenka, G.W. Heath, S. Inoue, S. Kahlmeier, P.T. Katzmarzyk, H.W. Kohl, E.V. Lambert, I.M. Lee, G. Leetongin, F. Lobelo, R.J.F. Loos, B. Marcus, B.W. Martin, N. Owen, D.C. Parra, M. Pratt, P. Puska, D. Ogilvie, R.S. Reis, J.F. Sallis, O.L. Sarmiento, J.C. Wells, Global physical activity levels: Surveillance progress, pitfalls, and prospects, *Lancet*. (2012). [https://doi.org/10.1016/S0140-6736\(12\)60646-1](https://doi.org/10.1016/S0140-6736(12)60646-1).
2. I.M. Lee, E.J. Shiroma, F. Lobelo, P. Puska, S.N. Blair, P.T. Katzmarzyk, J.R. Alkandari, L.B. Andersen, A.E. Bauman, R.C. Brownson, F.C. Bull, C.L. Craig, U. Ekelund, S. Goenka, R. Guthold, P.C. Hallal, W.L. Haskell, G.W. Heath, S. Inoue, S. Kahlmeier, H.W. Kohl, E.V. Lambert, G. Leetongin, R.J.F. Loos, B. Marcus, B.W. Martin, N. Owen, D.C. Parra, M. Pratt, D. Ogilvie, R.S. Reis, J.F. Sallis, O.L. Sarmiento, J.C. Wells, Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy, *Lancet*. (2012). [https://doi.org/10.1016/S0140-6736\(12\)61031-9](https://doi.org/10.1016/S0140-6736(12)61031-9).
3. J.E. Manson, P. Greenland, A.Z. LaCroix, M.L. Stefanick, C.P. Mouton, A. Oberman, M.G. Perri, D.S. Sheps, M.B. Pettinger, D.S. Siscovick, Walking Compared with Vigorous Exercise for the Prevention of Cardiovascular Events in Women, *N. Engl. J. Med.* 347 (2002) 716–725. <https://doi.org/10.1056/NEJMoa021067>.
4. K.Y. Wolin, Y. Yan, G.A. Colditz, I.M. Lee, Physical activity and colon cancer prevention: A meta-analysis, *Br. J. Cancer*. (2009). <https://doi.org/10.1038/sj.bjc.6604917>.
5. M. Duclos, J.M. Oppert, B. Verges, V. Coliche, J.F. Gautier, Y. Guezennec, G. Reach, G. Strauch, Physical activity and type 2 diabetes. Recommendations of the SFD (Francophone Diabetes Society) diabetes and physical activity working group, *Diabetes Metab.* 39 (2013) 205–216. <https://doi.org/10.1016/j.diabet.2013.03.005>.
6. B.K. Pedersen, B. Saltin, Exercise as medicine - Evidence for prescribing exercise as therapy in 26 different chronic diseases, *Scand. J. Med. Sci. Sport.* 25 (2015) 1–72. <https://doi.org/10.1111/sms.12581>.
7. U. Ekelund, J. Tarp, J. Steene-Johannessen, B.H. Hansen, B. Jefferis, M.W. Fagerland, P. Whincup, K.M. Diaz, S.P. Hooker, A. Chernofsky, M.G. Larson, N. Spartano, R.S. Vasan, I.M. Dohrn, M. Hagströmer, C. Edwardson, T. Yates, E. Shiroma, S.A. Anderssen, I.M. Lee, Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: Systematic review and harmonised meta-analysis, *BMJ*. 366 (2019) 1–10. <https://doi.org/10.1136/bmj.l4570>.
8. J.Y. Chau, A.C. Grunseit, T. Chey, E. Stamatakis, W.J. Brown, C.E. Matthews, A.E. Bauman, H.P. Van Der Ploeg, Daily sitting time and all-cause mortality: A meta-analysis, *PLoS One*. 8 (2013). <https://doi.org/10.1371/journal.pone.0080000>.
9. S. Deterding, D. Dixon, R. Khaled, L. Nacke, From game design elements to gamefulness: Defining “gamification,” *Proc. 15th Int. Acad. MindTrek Conf. Envisioning Futur. Media Environ. MindTrek 2011*. (2011) 9–15. <https://doi.org/10.1145/2181037.2181040>.
10. B. Cugelman, Gamification: What it is and why it matters to digital health behavior change developers, *J. Med. Internet Res.* 15 (2013) 1–6. <https://doi.org/10.2196/games.3139>.
11. J. Hamari, J. Koivisto, H. Sarsa, Does gamification work? - A literature review of empirical studies on gamification, *Proc. Annu. Hawaii Int. Conf. Syst. Sci.* (2014) 3025–3034. <https://doi.org/10.1109/HICSS.2014.377>.
12. D. Johnson, S. Deterding, K.A. Kuhn, A. Staneva, S. Stoyanov, L. Hides, Gamification for

- health and wellbeing: A systematic review of the literature, *Internet Interv.* 6 (2016) 89–106. <https://doi.org/10.1016/j.invent.2016.10.002>.
13. T. Munson, S., Poole, E., Perry, D. B., & Peyton, Gamification and health, in: *Gameful World Approaches, Issues Appl.*, 2014: pp. 597–624.
 14. J. Koivisto, J. Hamari, Gamification of physical activity: A systematic literature review of comparison studies, *CEUR Workshop Proc.* 2359 (2019) 106–117.
 15. K. Seaborn, D.I. Fels, Gamification in theory and action: A survey, *Int. J. Hum. Comput. Stud.* (2015). <https://doi.org/10.1016/j.ijhcs.2014.09.006>.
 16. D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, PRISMA 2009 Checklist, *Ann. Intern. Med.* (2014).
 17. D.S. Quintana, From pre-registration to publication: A non-technical primer for conducting a meta-analysis to synthesize correlational data, *Front. Psychol.* (2015). <https://doi.org/10.3389/fpsyg.2015.01549>.
 18. A. Mazéas, M. Duclos, B. Pereira, A. Chalabaev, Does gamification improve physical activity? A systematic review and meta-analysis. *OSF*, (2020). <https://doi.org/10.17605/OSF.IO/X2483>.
 19. D. Moreau, B. Gamble, Conducting a Meta-Analysis in the Age of Open Science: Tools, Tips, and Practical Recommendations, *Psychol. Methods.* (2020). <https://doi.org/10.1037/met0000351>.
 20. H.J.F. Helmerhorst, S. Brage, J. Warren, H. Besson, U. Ekelund, A systematic review of reliability and objective criterion-related validity of physical activity questionnaires, *Int. J. Behav. Nutr. Phys. Act.* (2012). <https://doi.org/10.1186/1479-5868-9-103>.
 21. W.V. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, *Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)*. Cochrane, 2019., Handbook. (2019).
 22. S.P. Hozo, B. Djulbegovic, I. Hozo, Estimating the mean and variance from the median, range, and the size of a sample, *BMC Med. Res. Methodol.* (2005). <https://doi.org/10.1186/1471-2288-5-13>.
 23. R. DerSimonian, N. Laird, Meta-analysis in clinical trials revisited, *Contemp. Clin. Trials.* (2015). <https://doi.org/10.1016/j.cct.2015.09.002>.
 24. J. Inthout, J.P. Ioannidis, G.F. Borm, The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method, *BMC Med. Res. Methodol.* (2014). <https://doi.org/10.1186/1471-2288-14-25>.
 25. L. V. Hedges, Distribution Theory for Glass's Estimator of Effect size and Related Estimators, *J. Educ. Stat.* (1981). <https://doi.org/10.3102/10769986006002107>.
 26. J. Cohen, *Statistical Power for the Behavioral Sciences* (2nd Edition), 1988.
 27. J.P.T. Higgins, S.G. Thompson, J.J. Deeks, D.G. Altman, Measuring inconsistency in meta-analyses Testing for heterogeneity, *BMJ.* (2003).
 28. B. Baujat, C. Mahé, J.P. Pignon, C. Hill, A graphical method for exploring heterogeneity in meta-analyses: Application to a meta-analysis of 65 trials, *Stat. Med.* (2002). <https://doi.org/10.1002/sim.1221>.
 29. M. Egger, G.D. Smith, M. Schneider, C. Minder, Bias in meta-analysis detected by a simple, graphical test, *Br. Med. J.* (1997). <https://doi.org/10.1136/bmj.315.7109.629>.
 30. S. Duval, R. Tweedie, Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis, *Biometrics.* (2000). <https://doi.org/10.1111/j.0006-341X.2000.00455.x>.
 31. M. Kossmeier, U.S. Tran, M. Voracek, Power-Enhanced Funnel Plots for Meta-Analysis: The Sunset Funnel Plot, *Zeitschrift Fur Psychol. / J. Psychol.* (2020). <https://doi.org/10.1027/2151-2604/a000392>.

32. M. Harrer, P. Cuijpers, T. Furukawa, D.D. Ebert, *Doing Meta-Analysis in R: A Hands-on Guide.*, Prot. Lab. (2019).
33. L.A. McGuinness, J.P.T. Higgins, Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments, in: *Res. Synth. Methods*, 2020. <https://doi.org/10.1002/jrsm.1411>.
34. A. Allam, Z. Kostova, K. Nakamoto, P.J. Schulz, The effect of social support features and gamification on a web-based intervention for rheumatoid arthritis patients: Randomized controlled trial, *J. Med. Internet Res.* 17 (2015) e14. <https://doi.org/10.2196/jmir.3510>.
35. R. Corepal, P. Best, R. O'Neill, F. Kee, J. Badham, L. Dunne, S. Miller, P. Connolly, M.E. Cupples, E.M.F. Van Sluijs, M.A. Tully, R.F. Hunter, A feasibility study of "The StepSmart Challenge" to promote physical activity in adolescents, *Pilot Feasibility Stud.* (2019). <https://doi.org/10.1186/s40814-019-0523-5>.
36. M.S. Haque, M. Kangas, T. Jämsä, A persuasive mhealth behavioral change intervention for promoting physical activity in the workplace: Feasibility randomized controlled trial, *J. Med. Internet Res.* (2020). <https://doi.org/10.2196/15083>.
37. C. Höchsmann, O. Müller, M. Ambühl, C. Klenk, K. Königstein, D. Infanger, S.P. Walz, A. Schmidt-Trucksäss, Novel Smartphone Game Improves Physical Activity Behavior in Type 2 Diabetes, *Am. J. Prev. Med.* (2019). <https://doi.org/10.1016/j.amepre.2019.02.017>.
38. G.W. Kurtzman, S.C. Day, D.S. Small, M. Lynch, J. Zhu, W. Wang, C.A.L. Rareshide, M.S. Patel, Social Incentives and Gamification to Promote Weight Loss: The LOSE IT Randomized, Controlled Trial, *J. Gen. Intern. Med.* (2018). <https://doi.org/10.1007/s11606-018-4552-1>.
39. A.-M. Leinonen, R. Pyky, R. Ahola, M. Kangas, P. Siirtola, T. Luoto, H. Enwald, T.M. Ikäheimo, J. Röning, S. Keinänen-Kiukaanniemi, M. Mäntysaari, R. Korpelainen, T. Jämsä, Feasibility of Gamified Mobile Service Aimed at Physical Activation in Young Men: Population-Based Randomized Controlled Study (MOPO), *JMIR MHealth UHealth*. (2017). <https://doi.org/10.2196/mhealth.6675>.
40. C. Maher, M. Ferguson, C. Vandelanotte, R. Plotnikoff, I. De Bourdeaudhuij, S. Thomas, K. Nelson-Field, T. Olds, A web-based, social networking physical activity intervention for insufficiently active adults delivered via Facebook app: Randomized controlled trial, *J. Med. Internet Res.* (2015). <https://doi.org/10.2196/jmir.4086>.
41. M. Nishiwaki, A. Kuriyama, Y. Ikegami, N. Nakashima, N. Matsumoto, A pilot crossover study: Effects of an intervention using an activity monitor with computerized game functions on physical activity and body composition, *J. Physiol. Anthropol.* (2014). <https://doi.org/10.1186/1880-6805-33-35>.
42. M.S. Patel, E.J. Benjamin, K.G. Volpp, C.S. Fox, D.S. Small, J.M. Massaro, J.J. Lee, V. Hilbert, M. Valentino, D.H. Taylor, E.S. Manders, K. Mutalik, J. Zhu, W. Wang, J.M. Murabito, Effect of a game-based intervention designed to enhance social incentives to increase physical activity among families: The BE FIT randomized clinical trial, *JAMA Intern. Med.* 177 (2017) 1586–1593. <https://doi.org/10.1001/jamainternmed.2017.3458>.
43. M.S. Patel, D.S. Small, J.D. Harrison, M.P. Fortunato, A.L. Oon, C.A.L. Rareshide, G. Reh, G. Schwartz, J. Guszcz, D. Steier, P. Kalra, V. Hilbert, Effectiveness of Behaviorally Designed Gamification Interventions with Social Incentives for Increasing Physical Activity among Overweight and Obese Adults Across the United States: The STEP UP Randomized Clinical Trial, *JAMA Intern. Med.* 179 (2019) 1624–1632. <https://doi.org/10.1001/jamainternmed.2019.3505>.
44. L. Paul, S. Wyke, S. Brewster, N. Sattar, J.M.R. Gill, G. Alexander, D. Rafferty, A.K. McFadyen, A. Ramsay, A. Dybus, Increasing physical activity in stroke survivors using STARFISH, an interactive mobile phone application: A pilot study, *Top. Stroke Rehabil.* (2016). <https://doi.org/10.1080/10749357.2015.1122266>.

45. K. Thorsteinsen, J. Vittersø, G.B. Svendsen, Increasing physical activity efficiently: An experimental pilot study of a website and mobile phone intervention, *Int. J. Telemed. Appl.* (2014). <https://doi.org/10.1155/2014/746232>.
46. K. Dadaczynski, S. Schieman, O. Backhaus, Promoting physical activity in worksite settings: Results of a German pilot study of the online intervention Healingo fit, *BMC Public Health.* (2017). <https://doi.org/10.1186/s12889-017-4697-6>.
47. O. Zuckerman, A. Gal-Oz, Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity, *Pers. Ubiquitous Comput.* (2014). <https://doi.org/10.1007/s00779-014-0783-2>.
48. A. Direito, Y. Jiang, R. Whittaker, R. Maddison, Apps for IMproving FITness and increasing physical activity among young people: The AIMFIT pragmatic randomized controlled trial, *J. Med. Internet Res.* (2015). <https://doi.org/10.2196/jmir.4568>.
49. S.M. Edney, T.S. Olds, J.C. Ryan, C. Vandelanotte, R.C. Plotnikoff, R.G. Curtis, C.A. Maher, A Social Networking and Gamified App to Increase Physical Activity: Cluster RCT, *Am. J. Prev. Med.* (2020). <https://doi.org/10.1016/j.amepre.2019.09.009>.
50. A. Garde, A. Umedaly, S.M. Abulnaga, L. Robertson, A. Junker, J.P. Chanoine, J.M. Ansermino, G.A. Dumont, Assessment of a Mobile Game ("MobileKids Monster Manor") to Promote Physical Activity among Children, *Games Health J.* (2015). <https://doi.org/10.1089/g4h.2014.0095>.
51. A. Garde, A. Umedaly, S.M. Abulnaga, A. Junker, J.P. Chanoine, M. Johnson, J.M. Ansermino, G.A. Dumont, Evaluation of a Novel Mobile Exergame in a School-Based Environment, *Cyberpsychology, Behav. Soc. Netw.* (2016). <https://doi.org/10.1089/cyber.2015.0281>.
52. A. Garde, M. Chowdhury, A.U. Rollinson, M. Johnson, P. Prescod, J.P. Chanoine, J.M. Ansermino, G.A. Dumont, A multi-week assessment of a mobile exergame intervention in an elementary school, *Games Health J.* (2018). <https://doi.org/10.1089/g4h.2017.0023>.
53. M. Gotsis, H. Wang, D. Spruijt-Metz, M. Jordan-Marsh, T.W. Valente, Wellness Partners: Design and evaluation of a web-based physical activity diary with social gaming features for adults, *J. Med. Internet Res.* (2013). <https://doi.org/10.2196/resprot.2132>.
54. A.L. Gremaud, L.J. Carr, J.E. Simmering, N.J. Evans, J.F. Cremer, A.M. Segre, L.A. Polgreen, P.M. Polgreen, Gamifying accelerometer use increases physical activity levels of sedentary office workers, *J. Am. Heart Assoc.* (2018). <https://doi.org/10.1161/JAHA.117.007735>.
55. R.G. Orwin, J.L. Vevea, Evaluating coding decisions, in: *Hand. Res. Synth. Meta-Analysis*, 2nd Ed., 2009.
56. J.O. Prochaska, C.C. Di Clemente, Transtheoretical therapy: Toward a more integrative model of change, *Psychotherapy.* (1982). <https://doi.org/10.1037/h0088437>.
57. A. Bandura, Health promotion from the perspective of social cognitive theory, *Psychol. Heal.* (1998). <https://doi.org/10.1080/08870449808407422>.
58. I. Ajzen, The theory of planned behavior, *Organ. Behav. Hum. Decis. Process.* (1991). [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T).
59. R. Schwarzer, Modeling health behavior change: How to predict and modify the adoption and maintenance of health behaviors, *Appl. Psychol.* (2008). <https://doi.org/10.1111/j.1464-0597.2007.00325.x>.
60. R.M. Ryan, E.L. Deci, Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being, *Am. Psychol.* (2000). <https://doi.org/10.1037/0003-066X.55.1.68>.
61. S. Gomersall, C. Maher, C. English, A. Rowlands, T. Olds, Time Regained : When People Stop a Physical Activity Program , How Does Their Time Use Change ? A Randomised Controlled Trial, (2015) 1–18. <https://doi.org/10.1371/journal.pone.0126665>.

62. I.M. Lee, E.J. Shiroma, M. Kamada, D.R. Bassett, C.E. Matthews, J.E. Buring, Association of Step Volume and Intensity with All-Cause Mortality in Older Women, *JAMA Intern. Med.* 179 (2019) 1105–1112. <https://doi.org/10.1001/jamainternmed.2019.0899>.
63. T. Yates, F. Zaccardi, N.N. Dhalwani, M.J. Davies, K. Bakrania, C.A. Celis-Morales, J.M.R. Gill, P.W. Franks, K. Khunti, Association of walking pace and handgrip strength with all-cause, cardiovascular, and cancer mortality: A UK Biobank observational study, *Eur. Heart J.* 38 (2017) 3232–3240. <https://doi.org/10.1093/eurheartj/ehx449>.
64. P. Oja, P. Kelly, E.M. Murtagh, M.H. Murphy, C. Foster, S. Titze, Effects of frequency, intensity, duration and volume of walking interventions on CVD risk factors: A systematic review and meta-regression analysis of randomised controlled trials among inactive healthy adults, *Br. J. Sports Med.* 52 (2018) 769–775. <https://doi.org/10.1136/bjsports-2017-098558>.
65. E.M. Murtagh, L. Nichols, M.A. Mohammed, R. Holder, A.M. Nevill, M.H. Murphy, The effect of walking on risk factors for cardiovascular disease: An updated systematic review and meta-analysis of randomised control trials, *Prev. Med. (Baltim).* 72 (2015) 34–43. <https://doi.org/10.1016/j.ypmed.2014.12.041>.
66. H.L. Gao, H.X. Gao, F.M. Sun, L. Zhang, Effects of walking on body composition in perimenopausal and postmenopausal women: A systematic review and meta-analysis, *Menopause.* 23 (2016) 928–934. <https://doi.org/10.1097/GME.0000000000000627>.
67. K.L. Piercy, R.P. Troiano, R.M. Ballard, S.A. Carlson, J.E. Fulton, D.A. Galuska, S.M. George, R.D. Olson, The physical activity guidelines for Americans, *JAMA - J. Am. Med. Assoc.* (2018). <https://doi.org/10.1001/jama.2018.14854>.
68. E.A. Finkelstein, B.A. Haaland, M. Bilger, A. Sahasranaman, R.A. Sloan, E.E.K. Nang, K.R. Evenson, Effectiveness of activity trackers with and without incentives to increase physical activity (TRIPPA): a randomised controlled trial, *Lancet Diabetes Endocrinol.* 4 (2016) 983–995. [https://doi.org/10.1016/S2213-8587\(16\)30284-4](https://doi.org/10.1016/S2213-8587(16)30284-4).
69. A. Romeo, S. Edney, R. Plotnikoff, R. Curtis, J. Ryan, I. Sanders, A. Crozier, C. Maher, Can Smartphone Apps Increase Physical Activity? Systematic Review and Meta-Analysis, *J. Med. Internet Res.* 21 (2019) e12053. <https://doi.org/10.2196/12053>.
70. X.S. Chen, S. Changolkar, A.S. Navathe, K.A. Linn, G. Reh, G. Schwartz, D. Steier, S. Godby, M. Balachandran, J.D. Harrison, C.A.L. Rareshide, M.S. Patel, Association between behavioral phenotypes and response to a physical activity intervention using gamification and social incentives: Secondary analysis of the STEP up randomized clinical trial, *PLoS One.* 15 (2020) 1–13. <https://doi.org/10.1371/journal.pone.0239288>.
71. R.E. Rhodes, I. Janssen, S.S.D. Bredin, D.E.R. Warburton, A. Bauman, Physical activity: Health impact, prevalence, correlates and interventions, *Psychol. Heal.* (2017). <https://doi.org/10.1080/08870446.2017.1325486>.
72. F. Faggiano, E. Allara, F. Giannotta, R. Molinar, H. Sumnall, R. Wiers, S. Michie, L. Collins, P. Conrod, Europe Needs a Central, Transparent, and Evidence-Based Approval Process for Behavioural Prevention Interventions, *PLoS Med.* (2014). <https://doi.org/10.1371/journal.pmed.1001740>.
73. C. Höchsmann, D. Infanger, C. Klenk, K. Königstein, S.P. Walz, A. Schmidt-Trucksäss, Effectiveness of a behavior change technique-based smartphone game to improve intrinsic motivation and physical activity adherence in patients with type 2 diabetes: Randomized controlled trial, *J. Med. Internet Res.* (2019). <https://doi.org/10.2196/11444>.
74. S. Nicholson, A User-Centered Theoretical Framework for Meaningful Gamification, *Games+ Learn. Soc.* (2012).
75. J. Jetten, C. Haslam, S.A. Haslam, G. Dingle, J.M. Jones, How groups affect our health and well-being: The path from theory to policy, *Soc. Issues Policy Rev.* 8 (2014) 103–130. <https://doi.org/10.1111/sipr.12003>.

Abbreviations

GRADE: grading of recommendations assessment, development, and evaluation

HAPA: health action process approach

LPA: light physical activity

MD: mean difference

MET: metabolic equivalent of task

MPA: moderate physical activity

MVPA: moderate-to-vigorous physical activity

OSF: open science framework

PA: physical activity

PRISMA: preferred reporting items for systematic reviews and meta-analyses

RCT: randomized controlled trial

SB: sedentary behaviors

SCT: socio-cognitive theory

SDT: self-determination theory

SMD: standardized mean difference

T2DM: type 2 diabetes

TPB: theory of planned behavior

TV: television

VPA: vigorous physical activity

Preprint
JMIR Publications

Supplementary Files

Multimedia Appendixes

Supplementary materials with additional figures and search strategy for all databases.

URL: <https://asset.jmir.pub/assets/57aeaa07e4dfd82862e87d4657f497f0.pdf>